

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 1, 2005, 16:25:44 ; Search time 43 Seconds  
(without alignments)

22.568 Million cell updates/sec

Title: US-09-993-366-1

Perfect score: 66

Sequence: 1 KLVPFATLHERL 13

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Maximum Match 0%

Listing first 100 summaries

Database :

- 1: /cgn2\_6/ptodata/1/iaa/5A\_COMB.psp.\*
- 2: /cgn2\_6/ptodata/1/iaa/5B\_COMB.psp.\*
- 3: /cgn2\_6/ptodata/1/iaa/6A\_COMB.psp.\*
- 4: /cgn2\_6/ptodata/1/iaa/6B\_COMB.psp.\*
- 5: /cgn2\_6/ptodata/1/iaa/6C\_COMB.psp.\*
- 6: /cgn2\_6/ptodata/1/iaa/backfiles1.psp.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	66	100.0	382	4	US-09-800-729-206
2	66	100.0	396	4	US-09-800-729-207
3	66	100.0	397	4	US-09-079-030-123
4	66	100.0	426	4	US-09-949-016-10854
5	38	57.6	243	4	US-09-543-681A-4422
6	38	57.6	492	4	US-09-489-039A-14208
7	38	57.6	850	4	US-09-949-016-11324
8	37	56.1	90	4	US-09-540-236-3521
9	37	56.1	127	4	US-08-311-731A-226
10	37	56.1	161	4	US-09-270-767-33980
11	37	56.1	161	4	US-09-270-767-49197
12	37	56.1	193	4	US-09-949-016-10446
13	37	56.1	302	4	US-09-902-540-10445
14	37	56.1	339	3	US-08-758-280-1
15	37	56.1	339	3	US-08-758-280-2
16	37	56.1	339	3	US-08-964-614A-1
17	37	56.1	339	3	US-08-964-614A-2
18	37	56.1	339	4	US-09-538-092-1274
19	37	56.1	436	4	US-09-724-623-94
20	37	56.1	436	4	US-09-902-540-16282
21	37	56.1	445	4	US-09-634-238-273
22	37	56.1	467	4	US-09-252-991A-25450
23	37	56.1	623	4	US-09-489-039A-11811
24	37	56.1	705	2	US-08-663-566A-19
25	37	56.1	705	2	US-08-023-610-19
26	37	56.1	705	2	US-08-288-065A-19
27	37	56.1	705	2	US-08-362-240A-19

28	56.1	705	5	PCT-US95-10245-19	Sequence 19, Appl
29	56.1	862	4	US-09-248-796A-16013	Sequence 16013, A
30	56.1	1051	4	US-09-252-991A-16989	Sequence 16989, A
31	56.1	1162	2	US-08-663-566A-15	Sequence 15, Appl
32	56.1	1162	2	US-08-023-610-15	Sequence 15, Appl
33	56.1	1162	2	US-08-288-065A-15	Sequence 15, Appl
34	56.1	1162	2	US-08-362-240A-15	Sequence 15, Appl
35	56.1	1162	5	PCT-US95-10245-15	Sequence 15, Appl
36	56.1	1265	3	US-09-347-878-5	Sequence 5, Appl
37	56.1	1265	3	US-09-347-878-7	Sequence 7, Appl
38	56.1	1265	4	US-08-980-326-2	Sequence 2, Appl
39	56.1	1265	4	US-08-980-326-74	Sequence 74, Appl
40	54.5	527	4	US-09-489-039A-14144	Sequence 14144, A
41	54.5	715	4	US-09-902-540-12939	Sequence 12939, A
42	54.5	739	3	US-08-510-646B-33	Sequence 33, Appl
43	53.0	141	4	US-09-270-767-35522	Sequence 35522, A
44	53.0	141	4	US-09-270-767-50739	Sequence 50739, A
45	53.0	143	3	US-09-134-001C-2941	Sequence 2941, Ap
46	53.0	143	4	US-09-710-279-54	Sequence 54, Appl
47	53.0	143	4	US-09-710-279-2672	Sequence 2672, Ap
48	53.0	199	4	US-09-248-796A-20525	Sequence 20525, A
49	53.0	377	4	US-09-248-796A-20227	Sequence 20227, A
50	53.0	473	4	US-08-985-343-6	Sequence 6, Appl
51	53.0	475	3	US-08-274-121B-4	Sequence 4, Appl
52	53.0	476	2	US-08-569-150A-3	Sequence 3, Appl
53	53.0	476	4	US-09-171-937C-2	Sequence 2, Appl
54	53.0	583	4	US-09-902-540-14460	Sequence 14460, A
55	53.0	669	4	US-09-107-532A-6532	Sequence 6532, Ap
56	53.0	676	4	US-09-902-540-10948	Sequence 10948, A
57	53.0	720	4	US-09-328-352-4435	Sequence 4435, A
58	51.5	1033	4	US-09-252-991A-20611	Sequence 20611, A
59	51.5	105	4	US-09-727-739B-36	Sequence 36, Appl
60	51.5	200	3	US-08-952-796-15	Sequence 15, Appl
61	51.5	203	4	US-09-902-540-10232	Sequence 10232, A
62	51.5	243	4	US-09-079-030-119	Sequence 119, App
63	51.5	256	1	US-08-594-808B-7	Sequence 7, Appl
64	51.5	267	1	US-07-959-946-3	Sequence 3, Appl
65	51.5	267	1	US-08-333-577-3	Sequence 3, Appl
66	51.5	267	3	US-08-952-796-2	Sequence 2, Appl
67	51.5	267	4	US-08-940-136-260	Sequence 260, App
68	51.5	267	4	US-09-919-039-27	Sequence 27, Appl
69	51.5	267	5	PCT-US92-08634-3	Sequence 3, Appl
70	51.5	382	4	US-09-492-709A-299	Sequence 299, App
71	51.5	682	4	US-09-270-767-45039	Sequence 45039, A
72	51.5	747	4	US-09-824-574-6	Sequence 6, Appl
73	51.5	753	4	US-09-949-016-10892	Sequence 10892, A
74	51.5	877	3	US-08-911-853-31	Sequence 31, Appl
75	51.5	877	3	US-09-479-409-31	Sequence 31, Appl
76	51.5	877	3	US-09-479-453-31	Sequence 31, Appl
77	51.5	1041	3	US-08-898-978-2	Sequence 2, Appl
78	51.5	1041	3	US-09-372-858-2	Sequence 8, Appl
79	51.5	1101	3	US-09-770-170-8	Sequence 4320, Ap
80	50.0	93	4	US-09-621-976-4320	Sequence 6, Appl
81	50.0	108	4	US-09-668-262A-6	Sequence 18025, A
82	50.0	133	4	US-09-252-991A-18025	Sequence 18025, A
83	50.0	136	4	US-09-270-767-47225	Sequence 74, Appl
84	50.0	151	3	US-08-679-493A-74	Sequence 5176, Ap
85	50.0	189	4	US-09-583-110-5176	Sequence 26395, A
86	50.0	203	4	US-09-252-991A-26395	Sequence 153, App
87	50.0	211	4	US-09-774-639-153	Sequence 42708, A
88	50.0	223	4	US-09-270-767-42708	Sequence 5068, Ap
89	50.0	231	4	US-09-107-433-5068	Sequence 5270, Ap
90	50.0	261	4	US-09-583-110-5270	Sequence 2780, Ap
91	50.0	276	4	US-09-107-433-2780	Sequence 6935, Ap
92	50.0	316	4	US-09-328-352-6935	Sequence 10982, A
93	50.0	386	4	US-09-902-540-10982	Sequence 208, App
94	50.0	391	4	US-09-800-729-208	Sequence 4714, Ap
95	50.0	399	4	US-09-107-532A-4714	Sequence 32313, A
96	50.0	400	4	US-09-252-991A-32313	Sequence 6906, Ap
97	50.0	414	4	US-09-949-016-6906	Sequence 30457, A
98	50.0	419	4	US-09-252-991A-30457	Sequence 16, Appl
99	50.0	421	4	US-09-668-262A-16	Sequence 14, Appl
100	50.0	423	4	US-09-668-262A-14	

## ALIGNMENTS

## RESULT 1

US-09-800-729-206  
; Sequence 206, Application US/09800729  
; Patent No. 6605592  
; GENERAL INFORMATION:  
; APPLICANT: Ni et al.  
; TITLE OF INVENTION: 32 Human secreted proteins  
; FILE REFERENCE: P2044P1  
; CURRENT APPLICATION NUMBER: US/09/800,729  
; CURRENT FILING DATE: 2001-03-08  
; PRIOR APPLICATION NUMBER: PCT/US00/26013  
; PRIOR FILING DATE: 2000-09-22  
; PRIOR APPLICATION NUMBER: 60/155,709  
; PRIOR FILING DATE: 1999-09-24  
; NUMBER OF SEQ ID NOS: 217  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 206  
; LENGTH: 382  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-800-729-206

Query Match 100.0%; Score 66; DB 4; Length 382;  
Best Local Similarity 100.0%; Pred. No. 0.00025;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KLVPFATELHERL 13  
Db |||||  
79 KLVPFATELHERL 91

## RESULT 2

US-09-800-729-207  
; Sequence 207, Application US/09800729  
; Patent No. 6605592  
; GENERAL INFORMATION:  
; APPLICANT: Ni et al.  
; TITLE OF INVENTION: 32 Human secreted proteins  
; FILE REFERENCE: P2044P1  
; CURRENT APPLICATION NUMBER: US/09/800,729  
; CURRENT FILING DATE: 2001-03-08  
; PRIOR APPLICATION NUMBER: PCT/US00/26013  
; PRIOR FILING DATE: 2000-09-22  
; PRIOR APPLICATION NUMBER: 60/155,709  
; PRIOR FILING DATE: 1999-09-24  
; NUMBER OF SEQ ID NOS: 217  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 207  
; LENGTH: 396  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-800-729-207

Query Match 100.0%; Score 66; DB 4; Length 396;  
Best Local Similarity 100.0%; Pred. No. 0.00026;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KLVPFATELHERL 13  
Db |||||  
79 KLVPFATELHERL 91

## RESULT 3

US-09-079-030-123  
; Sequence 123, Application US/09079030  
; Patent No. 6635623  
; GENERAL INFORMATION:  
; APPLICANT: Guevera, Jz., Juan G.

APPLICANT: Hoogveen, Ron C.  
APPLICANT: Moore, Paul J.  
TITLE OF INVENTION: LIPOPROTEINS AS NUCLEIC ACID DELIVERY  
TITLE OF INVENTION: VECTORS FOR TRANSFECTION OF EUKARYOTIC CELLS  
NUMBER OF SEQUENCES: 229  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Arnold, White & Durkee  
STREET: P.O. Box 4433  
CITY: Houston  
STATE: Texas  
COUNTRY: USA  
ZIP: 77210  
COMPUTER READABLE FORM: disk  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/079,030  
FILING DATE: Concurrently Herewith  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: McMillian, Nabesla R.  
REGISTRATION NUMBER: P-43,363  
REFERENCE/DOCKET NUMBER: ARAG:003  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 512/418-3000  
TELEFAX: 512/474-7577  
INFORMATION FOR SEQ ID NO: 123:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 397 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
US-09-079-030-123

Query Match 100.0%; Score 66; DB 4; Length 397;  
Best Local Similarity 100.0%; Pred. No. 0.00026;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KLVPFATELHERL 13  
Db |||||  
30 KLVPFATELHERL 42

## RESULT 4

US-09-949-016-10854  
; Sequence 10854, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 10854  
; LENGTH: 426  
; TYPE: PRT  
; ORGANISM: Human  
US-09-949-016-10854

Query Match 100.0%; Score 66; DB 4; Length 426;  
Best Local Similarity 100.0%; Pred. No. 0.00028;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLVPFATELHERL 13  
Db 109 KLVPFATELHERL 121

## RESULT 5

US-09-543-681A-4422  
; Sequence 4422, Application US/09543681A  
; Patent No. 6605709  
; GENERAL INFORMATION:  
; APPLICANT: GARY BRETON  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS  
; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 2709.1002-001  
; CURRENT APPLICATION NUMBER: US/09/543.681A  
; CURRENT FILING DATE: 2000-04-05  
; PRIOR APPLICATION NUMBER: US 60/128,706  
; PRIOR FILING DATE: 1999-04-09  
; NUMBER OF SEQ ID NOS: 8344  
; SEQ ID NO 4422  
; LENGTH: 243  
; TYPE: PRT  
; ORGANISM: Proteus mirabilis  
US-09-543-681A-4422

Query Match 57.6%; Score 38; DB 4; Length 243;  
Best Local Similarity 87.5%; Pred. No. 26;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 4 PFATELHE 11  
Db 59 PFATELHE 66

## RESULT 6

US-09-489-039A-14208  
; Sequence 14208, Application US/09489039A  
; Patent No. 6610836  
; GENERAL INFORMATION:  
; APPLICANT: Gary Breton et. al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA  
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 2709.2004001  
; CURRENT APPLICATION NUMBER: US/09/489.039A  
; CURRENT FILING DATE: 2000-01-27  
; PRIOR APPLICATION NUMBER: US 60/117,747  
; PRIOR FILING DATE: 1999-01-29  
; NUMBER OF SEQ ID NOS: 14342  
; SEQ ID NO 14208  
; LENGTH: 492  
; TYPE: PRT  
; ORGANISM: Klebsiella pneumoniae  
US-09-489-039A-14208

Query Match 57.6%; Score 38; DB 4; Length 492;  
Best Local Similarity 63.6%; Pred. No. 56;  
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 2 LVPPFATELHER 12  
Db 152 LVPPFATELHER 162

## RESULT 7

US-09-949-016-11324  
; Sequence 11324, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307

; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 11324  
; LENGTH: 850  
; TYPE: PRT  
; ORGANISM: Human  
US-09-949-016-11324

Query Match 57.6%; Score 38; DB 4; Length 850;  
Best Local Similarity 72.7%; Pred. No. 1e+02;  
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KLVPFATELHE 11  
Db 373 KLVPFATELHE 383

## RESULT 8

US-09-540-236-3521  
; Sequence 3521, Application US/09540236  
; Patent No. 6673910  
; GENERAL INFORMATION:  
; APPLICANT: Gary L. Breton et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CAT  
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 2709.2005-001  
; CURRENT APPLICATION NUMBER: US/09/540,236  
; CURRENT FILING DATE: 2000-04-04  
; NUMBER OF SEQ ID NOS: 3840  
; SEQ ID NO 3521  
; LENGTH: 90  
; TYPE: PRT  
; ORGANISM: M.catarrhalis  
US-09-540-236-3521

Query Match 56.1%; Score 37; DB 4; Length 90;  
Best Local Similarity 77.8%; Pred. No. 13;  
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 4 PFATELHER 12  
Db 22 PFATELHER 30

## RESULT 9

US-08-311-731A-226  
; Sequence 226, Application US/08311731A  
; Patent No. 6583266  
; GENERAL INFORMATION:  
; APPLICANT: SMITH, DOUGLAS  
; APPLICANT: MAO, JEN-I  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES  
; TITLE OF INVENTION: RELATING TO MYCOBACTERIUM TUBERCULOSIS AND LAPRAE FOR  
; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS  
; NUMBER OF SEQUENCES: 411  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: WOLF, GREENFIELD & SACKS, P.C.  
; STREET: 600 ATLANTIC AVENUE  
; CITY: BOSTON  
; STATE: MASSACHUSETTS  
; COUNTRY: USA  
; ZIP: 02210  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible

```
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/311,731A
; FILING DATE:
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: GATES, EDWARD R.
; REGISTRATION NUMBER: 31,616
; REFERENCE/DOCKET NUMBER: C0044/7125
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/720-3500
; TELEFAX: 617/720-2441
; INFORMATION FOR SEQ ID NO: 226:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 127 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Mycobacterium leprae
; US-08-311-731A-226

Query Match 56.1%; Score 37; DB 4; Length 127;
Best Local Similarity 54.5%; Pred. No. 19;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 3 VPFATLHERL 13
Db 33 VPEAVSVHRI 43

RESULT 10
US-09-270-767-33980
; Sequence 33980, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 33980
; LENGTH: 161
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-33980

Query Match 56.1%; Score 37; DB 4; Length 161;
Best Local Similarity 60.0%; Pred. No. 25;
Matches 6; Conservative 3; Mismatches 3; Indels 1; Gaps 0;

QY 2 LVPPATELHE 11
Db 99 LEPYSTELHQ 108

RESULT 11
US-09-270-767-49197
; Sequence 49197, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517

; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 49197
; LENGTH: 161
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-49197

Query Match 56.1%; Score 37; DB 4; Length 161;
Best Local Similarity 60.0%; Pred. No. 25;
Matches 6; Conservative 3; Mismatches 3; Indels 1; Gaps 0;

QY 2 LVPPATELHE 11
Db 99 LEPYSTELHQ 108

RESULT 12
US-09-949-016-10446
; Sequence 10446, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10446
; LENGTH: 193
; TYPE: PRT
; ORGANISM: Human
; US-09-949-016-10446

Query Match 56.1%; Score 37; DB 4; Length 193;
Best Local Similarity 56.2%; Pred. No. 31;
Matches 9; Conservative 1; Mismatches 2; Indels 4; Gaps 1;

QY 1 KLVPPF----ATELHER 12
Db 96 KLVPMGFTTATEFHQR 111

RESULT 13
US-09-902-540-10445
; Sequence 10445, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 10445
; LENGTH: 302
; TYPE: PRT
; ORGANISM: Myxococcus xanthus
; US-09-902-540-10445
```

```

Query Match      56.1%; Score 37; DB 4; Length 302;
Best Local Similarity 66.7%; Pred. No. 50;
Matches      6; Conservative      2; Mismatches      1; Indels      0; Gaps      0;

QY      2 LVPPATELH 10
      |.:|:|:|
Db      226 LIPLATQLH 234

RESULT 14
US-08-758-280-1
; Sequence 1, Application US/08758280
; Patent No. 6037125
; GENERAL INFORMATION:
; APPLICANT: Hasty, Paul
; TITLE OF INVENTION: DISRUPTION OF THE MAMMALIAN Rad51
; TITLE OF INVENTION: PROTEIN AND DISRUPTION OF PROTEINS
; TITLE OF INVENTION: MAMMALIAN Rad51 FOR HINDERING CELL PROLIFERATION AND/OR
; TITLE OF INVENTION: VIABILITY OF PROLIFERATING CELLS
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PENNIE & EDMONDS
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/758,280
; FILING DATE: concurrently herewith
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Halluin, Albert P.
; REGISTRATION NUMBER: 25,227
; REFERENCE/DOCKET NUMBER: 8535-013-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 854-3660
; TELEFAX: (415) 854-3694
; TELEX: 66141
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 339 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; US-08-758-280-1

Query Match      56.1%; Score 37; DB 3; Length 339;
Best Local Similarity 56.2%; Pred. No. 57;
Matches      9; Conservative      1; Mismatches      2; Indels      4; Gaps      1;

QY      1 KLVPF---ATELHER 12
      |||||
Db      80 KLVPMGFTTATEFHQR 95

RESULT 15
US-08-758-280-2
; Sequence 2, Application US/08758280
; Patent No. 6037125
; GENERAL INFORMATION:
; APPLICANT: Hasty, Paul
; TITLE OF INVENTION: DISRUPTION OF THE MAMMALIAN Rad51
; TITLE OF INVENTION: PROTEIN AND DISRUPTION OF PROTEINS
; TITLE OF INVENTION: MAMMALIAN Rad51 FOR HINDERING CELL PROLIFERATION AND/OR

```

```

US-09-724-623-94
; ORGANISM: Lactobacillus rhamnosus
;
Query Match          56.1%; Score 37; DB 4; Length 436;
Best Local Similarity 77.8%; Pred. No. 75;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QV 5 FATELHERL 13

```

Db 380 FVTRLHERL 388  
||| |||||

## RESULT 20

US-09-902-540-16282  
; Sequence 16282, Application US/09902540  
; Patent No. 6833447  
; GENERAL INFORMATION:  
; APPLICANT: Goldman, Barry S.  
; APPLICANT: Hinkle, Gregory J.  
; APPLICANT: Slater, Steven C.  
; APPLICANT: Wiegand, Roger C.  
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof  
; FILE REFERENCE: 38-10(15849)B  
; CURRENT APPLICATION NUMBER: US/09/902,540  
; CURRENT FILING DATE: 2001-07-10  
; PRIOR APPLICATION NUMBER: 60/217,883  
; PRIOR FILING DATE: 2000-07-10  
; NUMBER OF SEQ ID NOS: 16825  
; SEQ ID NO 16282  
; LENGTH: 436  
; TYPE: PRT  
; ORGANISM: Myxococcus xanthus  
US-09-902-540-16282

Query Match 56.1%; Score 37; DB 4; Length 436;  
Best Local Similarity 60.0%; Pred. No. 75;  
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 4 PFATELHERL 13  
||| |||||  
Db 283 PFPELHDAI 292

## RESULT 21

US-09-634-238-273  
; Sequence 273, Application US/09634238  
; Patent No. 6544772  
; GENERAL INFORMATION:  
; APPLICANT: Glenn, Matthew  
; APPLICANT: Havukkala, Ilkka J.  
; APPLICANT: Bloksberg, Leonard, N.  
; APPLICANT: Lubbers, Mark W.  
; APPLICANT: Dekker, James  
; APPLICANT: Christenson, Anna C.  
; APPLICANT: Holland, Ross  
; APPLICANT: O'Toole, Paul W.  
; APPLICANT: Reid, Julian R.  
; APPLICANT: Coolbear, Timothy  
; TITLE OF INVENTION: Polynucleotides, materials incorporating  
; TITLE OF INVENTION: them and methods for using them.  
; FILE REFERENCE: 11000.1043U1  
; CURRENT APPLICATION NUMBER: US/09/634,238  
; CURRENT FILING DATE: 2000-08-08  
; NUMBER OF SEQ ID NOS: 422  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 273  
; LENGTH: 445  
; TYPE: PRT  
; ORGANISM: Lactobacillus rhamnosus  
US-09-634-238-273

Query Match 56.1%; Score 37; DB 4; Length 445;  
Best Local Similarity 77.8%; Pred. No. 77;  
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 5 FATELHERL 13  
||| |||||  
Db 387 FVTRLHERL 395

## RESULT 22

US-09-252-991A-25450  
; Sequence 25450, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 25450  
; LENGTH: 467  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-25450

Query Match 56.1%; Score 37; DB 4; Length 467;  
Best Local Similarity 53.8%; Pred. No. 81;  
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 1 KLVPPFATELHERL 13  
||| |||||  
Db 60 KVMKHAELHERI 72

## RESULT 23

US-09-489-039A-11811  
; Sequence 11811, Application US/09489039A  
; Patent No. 6610836  
; GENERAL INFORMATION:  
; APPLICANT: Gary Breton et. al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA  
; FILE REFERENCE: 2709.2004001  
; CURRENT APPLICATION NUMBER: US/09/489,039A  
; CURRENT FILING DATE: 2000-01-27  
; PRIOR APPLICATION NUMBER: US 60/117,747  
; PRIOR FILING DATE: 1999-01-29  
; NUMBER OF SEQ ID NOS: 14342  
; SEQ ID NO 11811  
; LENGTH: 623  
; TYPE: PRT  
; ORGANISM: Klebsiella pneumoniae  
US-09-489-039A-11811

Query Match 56.1%; Score 37; DB 4; Length 623;  
Best Local Similarity 54.5%; Pred. No. 11e+02;  
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 3 VPFATELHERL 13  
||| |||||  
Db 366 VPEGVEVHRV 376

## RESULT 24

US-08-663-566A-19  
; Sequence 19, Application US/08663566A  
; Patent No. 5853733  
; GENERAL INFORMATION:  
; APPLICANT: Cochran, Mark D  
; APPLICANT: Macdonald, Richard D  
; TITLE OF INVENTION: Recombinant Herpesvirus of Turkeys  
; TITLE OF INVENTION: and Uses Thereof  
; NUMBER OF SEQUENCES: 56  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: John P. White  
; STREET: 1185 Avenue of the Americas  
; CITY: New York

```
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/663,566A
; FILING DATE: June 13, 1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P
; REGISTRATION NUMBER: 28,678
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)278-0400
; TELEFAX: (212)391-0526
; TELEX: 422523
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 705 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; US-08-663-566A-19

Query Match 56.1%; Score 37; DB 2; Length 705;
Best Local Similarity 54.5%; Pred. No. 1.3e+02;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 3 VPPATELHERL 13
Db 316 IPPATOLQARI 326

RESULT 25
US-08-023-610-19
; Sequence 19, Application US/08023610
; Patent No. 5928648
; GENERAL INFORMATION:
; APPLICANT: Cochran Ph.D., Mark D
; APPLICANT: Macdonald Ph.D., Richard D
; TITLE OF INVENTION: Recombinant Herpesvirus of Turkeys
; TITLE OF INVENTION: and Uses Thereof
; NUMBER OF SEQUENCES: 47
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: John P. White
; STREET: 30 Rockefeller Plaza
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10112
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/023,610
; FILING DATE: February 26, 1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White Esq, John P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)977-9550
; TELEFAX: (212)664-0525
; TELEX: 422523
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 705 amino acids
```

```
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; US-08-023-610-19

Query Match 56.1%; Score 37; DB 2; Length 705;
Best Local Similarity 54.5%; Pred. No. 1.3e+02;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 3 VPPATELHERL 13
Db 316 IPPATOLQARI 326

RESULT 26
US-08-288-065A-19
; Sequence 19, Application US/08288065A
; Patent No. 5961982
; GENERAL INFORMATION:
; APPLICANT: Cochran, Mark D
; APPLICANT: Macdonald, Richard D
; TITLE OF INVENTION: Recombinant Herpesvirus of Turkeys S-
; TITLE OF INVENTION: HVT-050 and Uses Thereof
; NUMBER OF SEQUENCES: 56
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: John P. White
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/288,065A
; FILING DATE: Aug-09-94
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)278-0400
; TELEFAX: (212)391-0526
; TELEX: 422523
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 705 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; US-08-288-065A-19

Query Match 56.1%; Score 37; DB 2; Length 705;
Best Local Similarity 54.5%; Pred. No. 1.3e+02;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 3 VPPATELHERL 13
Db 316 IPPATOLQARI 326

RESULT 27
US-08-362-240A-19
; Sequence 19, Application US/08362240A
; Patent No. 5965138
; GENERAL INFORMATION:
; APPLICANT: Cochran, Mark D
```



APPLICANT: Junker, David  
APPLICANT: Wild, Martha A  
TITLE OF INVENTION: Recombinant Herpesvirus and Uses Thereof  
NUMBER OF SEQUENCES: 60  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: John P. White  
STREET: 1185 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA: US/08/362,240A  
APPLICATION NUMBER: US/08/362,240A  
FILING DATE: Dec-22-94  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: White, John P  
REGISTRATION NUMBER: 28,678  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212)278-0400  
TELEFAX: (212)391-0526  
TELEX: 422523  
INFORMATION FOR SEQ ID NO: 19:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 705 amino acids  
TYPE: amino acid  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
US-08-362-240A-19

Query Match 56.1%; Score 37; DB 2; Length 705;  
Best Local Similarity 54.5%; Pred. No. 1.3e+02;  
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 3 VPFPATELHERL 13  
Db 316 IPFATQLQARI 326

RESULT 28  
PCT-US95-10245-19  
Sequence 19, Application PC/TUS9510245  
GENERAL INFORMATION:  
APPLICANT: SYNTRO CORPORATION  
TITLE OF INVENTION: Recombinant Herpesvirus of Turkeys And Uses Thereof  
NUMBER OF SEQUENCES: 60  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: John P. White  
STREET: 1185 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/10245  
FILING DATE: 09-AUG-1995  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: White, John P  
REGISTRATION NUMBER: 28,678  
TELECOMMUNICATION INFORMATION:

TELEPHONE: (212)278-0400  
TELEFAX: (212)391-0526  
TELEX: 422523  
INFORMATION FOR SEQ ID NO: 19:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 705 amino acids  
TYPE: amino acid  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
PCT-US95-10245-19

Query Match 56.1%; Score 37; DB 5; Length 705;  
Best Local Similarity 54.5%; Pred. No. 1.3e+02;  
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 3 VPFPATELHERL 13  
Db 316 IPFATQLQARI 326

RESULT 29  
US-09-248-796A-16013  
Sequence 16013, Application US/09248796A  
Patent No. 6747137  
GENERAL INFORMATION:  
APPLICANT: Keith Weinstock et al  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICA  
TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS  
FILE REFERENCE: 107196.132  
CURRENT APPLICATION NUMBER: US/09/248,796A  
CURRENT FILING DATE: 1999-02-12  
PRIOR APPLICATION NUMBER: US 60/074,725  
PRIOR FILING DATE: 1998-02-13  
PRIOR APPLICATION NUMBER: US 60/096,409  
PRIOR FILING DATE: 1998-08-13  
NUMBER OF SEQ ID NOS: 28208  
SEQ ID NO 16013  
LENGTH: 862  
TYPE: PRT  
ORGANISM: Candida albicans  
US-09-248-796A-16013

Query Match 56.1%; Score 37; DB 4; Length 862;  
Best Local Similarity 70.0%; Pred. No. 1.6e+02;  
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 2 LVFPATELHE 11  
Db 60 LVEFLTEIHE 69

RESULT 30  
US-09-252-991A-16989  
Sequence 16989, Application US/09252991A  
Patent No. 6551795  
GENERAL INFORMATION:  
APPLICANT: Marc J. Rubenfield et al.  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
FILE REFERENCE: 107196.136  
CURRENT APPLICATION NUMBER: US/09/252,991A  
CURRENT FILING DATE: 1999-02-18  
PRIOR APPLICATION NUMBER: US 60/074,788  
PRIOR FILING DATE: 1998-02-18  
PRIOR APPLICATION NUMBER: US 60/094,190  
PRIOR FILING DATE: 1998-07-27  
NUMBER OF SEQ ID NOS: 33142  
SEQ ID NO 16989  
LENGTH: 1051  
TYPE: PRT  
ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-16989

Query Match 56.1%; Score 37; DB 4; Length 1051;  
 Best Local Similarity 63.6%; Pred. No. 2e+02;  
 Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 LVPFATELHER 12  
 :|||||:  
 Db 949 IVEFAKELHEQ 959

RESULT 31

US-08-663-566A-15  
 ; Sequence 15, Application US/08663566A

; Patent No 5853733

; GENERAL INFORMATION:

; APPLICANT: Cochran, Mark D

; APPLICANT: Macdonald, Richard D

; TITLE OF INVENTION: Recombinant Herpesvirus of Turkeys

; TITLE OF INVENTION: and Uses Thereof

; NUMBER OF SEQUENCES: 56

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: John P. White

; STREET: 1185 Avenue of the Americas

; CITY: New York

; STATE: New York

; COUNTRY: USA

; ZIP: 10036

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/663,566A

; FILING DATE: June 13, 1996

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: White, John P

; REGISTRATION NUMBER: 28,678

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (212)278-0400

; TELEFAX: (212)391-0526

; TELEX: 422523

; INFORMATION FOR SEQ ID NO: 15:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1162 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-08-663-566A-15

Query Match 56.1%; Score 37; DB 2; Length 1162;  
 Best Local Similarity 54.5%; Pred. No. 2.2e+02;  
 Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 3 VPFATELHERL 13  
 :|||||:  
 Db 769 IFFATQLQARI 779

RESULT 32

US-08-023-610-15

; Sequence 15, Application US/08023610

; Patent No. 5928648

; GENERAL INFORMATION:

; APPLICANT: Cochran Ph.D, Mark D

; APPLICANT: Macdonald Ph.D., Richard D

; TITLE OF INVENTION: Recombinant Herpesvirus of Turkeys

; TITLE OF INVENTION: and Uses Thereof

; NUMBER OF SEQUENCES: 47

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: John P. White

; STREET: 30 Rockefeller Plaza  
 ; CITY: New York  
 ; STATE: New York  
 ; COUNTRY: USA  
 ; ZIP: 10112  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/023,610  
 ; FILING DATE: February 26, 1993  
 ; CLASSIFICATION: 435  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: White Esq, John P  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (212)977-9550  
 ; TELEFAX: (212)664-0525  
 ; TELEX: 422523  
 ; INFORMATION FOR SEQ ID NO: 15:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 1162 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein  
 ; US-08-023-610-15

Query Match 56.1%; Score 37; DB 2; Length 1162;  
 Best Local Similarity 54.5%; Pred. No. 2.2e+02;  
 Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 3 VPFATELHERL 13  
 :|||||:  
 Db 769 IFFATQLQARI 779

RESULT 33

US-08-288-065A-15

; Sequence 15, Application US/08288065A

; Patent No. 5961982

; GENERAL INFORMATION:

; APPLICANT: Cochran, Mark D

; APPLICANT: Macdonald, Richard D

; TITLE OF INVENTION: Recombinant Herpesvirus of Turkeys S-

; TITLE OF INVENTION: HVT-050 and Uses Thereof

; NUMBER OF SEQUENCES: 56

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: John P. White

; STREET: 1185 Avenue of the Americas

; CITY: New York

; STATE: New York

; COUNTRY: USA

; ZIP: 10036

; COMPUTER READABLE FORM: disk

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/288,065A

; FILING DATE: Aug-09-94

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: White, John P

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (212)278-0400

; TELEFAX: (212)391-0526

; TELEX: 422523

; INFORMATION FOR SEQ ID NO: 15:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1162 amino acids

; TYPE: amino acid

; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-288-065A-15

Query Match 56.1%; Score 37; DB 2; Length 1162;  
Best Local Similarity 54.5%; Pred. No. 2.2e+02;  
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 3 VPFAELHERL 13  
Db 769 IPFATQIARI 779

## RESULT 34

US-08-362-240A-15  
; Sequence 15, Application US/08362240A  
; Patent No. 5965138  
; GENERAL INFORMATION:  
; APPLICANT: Cochran, Mark D  
; APPLICANT: Junker, David  
; APPLICANT: Wild, Martha A  
; TITLE OF INVENTION: Recombinant Herpesvirus and Uses Thereof  
; NUMBER OF SEQUENCES: 60  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: John P. White  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/362,240A  
; FILING DATE: Dec-22-94  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: White, John P  
; REGISTRATION NUMBER: 28,678  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212)278-0400  
; TELEFAX: (212)391-0526  
; TELEX: 422523  
; INFORMATION FOR SEQ ID NO: 15:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1162 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-362-240A-15

Query Match 56.1%; Score 37; DB 2; Length 1162;  
Best Local Similarity 54.5%; Pred. No. 2.2e+02;  
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 3 VPFAELHERL 13  
Db 769 IPFATQIARI 779

## RESULT 35

PCT-US95-10245-15  
; Sequence 15, Application PC/TUS9510245  
; GENERAL INFORMATION:  
; APPLICANT: SYNTRO CORPORATION  
; TITLE OF INVENTION: Recombinant Herpesvirus of Turkeys And Uses Thereof  
; NUMBER OF SEQUENCES: 60  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: John P. White  
; STREET: 1185 Avenue of the Americas

; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US95/10245  
; FILING DATE: 09-AUG-1995  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: White, John P  
; REGISTRATION NUMBER: 28,678  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212)278-0400  
; TELEFAX: (212)391-0526  
; TELEX: 422523  
; INFORMATION FOR SEQ ID NO: 15:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1162 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; PCT-US95-10245-15

Query Match 56.1%; Score 37; DB 5; Length 1162;  
Best Local Similarity 54.5%; Pred. No. 2.2e+02;  
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 3 VPFAELHERL 13  
Db 769 IPFATQIARI 779

## RESULT 36

US-09-347-878-5  
; Sequence 5, Application US/09347878C  
; Patent No. 6376210  
; GENERAL INFORMATION:  
; APPLICANT: Yuan, Chong  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR ASSAYING ANALYTES  
; FILE REFERENCE: 25885-1651  
; CURRENT APPLICATION NUMBER: US/09/347,878C  
; CURRENT FILING DATE: 1999-07-06  
; NUMBER OF SEQ ID NOS: 75  
; SOFTWARE: Patent In Ver. 2.0  
; SEQ ID NO 5  
; LENGTH: 1265  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-347-878-5

Query Match 56.1%; Score 37; DB 3; Length 1265;  
Best Local Similarity 77.8%; Pred. No. 2.4e+02;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 5 FATELHERL 13  
Db 1137 FABELHERV 1145

## RESULT 37

US-09-347-878-7  
; Sequence 7, Application US/09347878C  
; Patent No. 6376210  
; GENERAL INFORMATION:  
; APPLICANT: Yuan, Chong  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR ASSAYING ANALYTES  
; FILE REFERENCE: 25885-1651  
; CURRENT APPLICATION NUMBER: US/09/347,878C

; CURRENT FILING DATE: 1999-07-06  
; NUMBER OF SEQ ID NOS: 75  
; SOFTWARE: Patent In Ver. 2.0  
; SEQ ID NO 7  
; LENGTH: 1265  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-347-878-7

Query Match 56.1%; Score 37; DB 3; Length 1265;  
Best Local Similarity 77.8%; Pred. No. 2.4e+02;  
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 5 FATELHERL 13  
||| |||||  
Db 1137 FAEELHERV 1145

RESULT 38  
US-08-980-326-2  
; Sequence 2, Application US/08980326  
; Patent No. 6703197  
; GENERAL INFORMATION:  
; APPLICANT: Gravel, Roy A.  
; APPLICANT: Rozen, Rima  
; APPLICANT: LeClerc, Daniel  
; APPLICANT: Goyette, Philippe  
; APPLICANT: Campeau, Eric  
; TITLE OF INVENTION: HUMAN METHIONINE SYNTHASE: CLONING, AND  
; TITLE OF INVENTION: METHODS FOR EVALUATING RISK OF NEURAL TUBE DEFECTS,  
; TITLE OF INVENTION: CARDIOVASCULAR DISEASE AND CANCER  
; FILE REFERENCE: 50004/002003  
; CURRENT APPLICATION NUMBER: US/08/980,326  
; CURRENT FILING DATE: 1997-11-26  
; EARLIER APPLICATION NUMBER: 60/050,310  
; EARLIER FILING DATE: 1997-06-20  
; EARLIER APPLICATION NUMBER: 60/031,964  
; EARLIER FILING DATE: 1996-11-27  
; NUMBER OF SEQ ID NOS: 75  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 2  
; LENGTH: 1265  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; NAME/KEY: VARIANT  
; LOCATION: (1)...(1265)  
; OTHER INFORMATION: Wild type methionine synthase polypeptide.  
US-08-980-326-2

Query Match 56.1%; Score 37; DB 4; Length 1265;  
Best Local Similarity 77.8%; Pred. No. 2.4e+02;  
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 5 FATELHERL 13  
||| |||||  
Db 1137 FAEELHERV 1145

RESULT 39  
US-08-980-326-74  
; Sequence 74, Application US/08980326  
; Patent No. 6703197  
; GENERAL INFORMATION:  
; APPLICANT: Gravel, Roy A.  
; APPLICANT: Rozen, Rima  
; APPLICANT: LeClerc, Daniel  
; APPLICANT: Goyette, Philippe  
; APPLICANT: Campeau, Eric  
; TITLE OF INVENTION: HUMAN METHIONINE SYNTHASE: CLONING, AND  
; TITLE OF INVENTION: METHODS FOR EVALUATING RISK OF NEURAL TUBE DEFECTS,  
; TITLE OF INVENTION: CARDIOVASCULAR DISEASE AND CANCER  
; FILE REFERENCE: 50004/002003

; CURRENT APPLICATION NUMBER: US/08/980,326  
; CURRENT FILING DATE: 1997-11-26  
; EARLIER APPLICATION NUMBER: 60/050,310  
; EARLIER FILING DATE: 1997-06-20  
; EARLIER APPLICATION NUMBER: 60/031,964  
; EARLIER FILING DATE: 1996-11-27  
; NUMBER OF SEQ ID NOS: 75  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 74  
; LENGTH: 1265  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: VARIANT  
; LOCATION: (1)...(1265)  
; OTHER INFORMATION: Mutant methionine synthases: 881  
; OTHER INFORMATION: can be Ile or deleted; 919 can be Asp or Gly; 920  
; OTHER INFORMATION: can be His or Asp.  
US-08-980-326-74

Query Match 56.1%; Score 37; DB 4; Length 1265;  
Best Local Similarity 77.8%; Pred. No. 2.4e+02;  
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 5 FATELHERL 13  
||| |||||  
Db 1137 FAEELHERV 1145

RESULT 40  
US-09-489-039A-14144  
; Sequence 14144, Application US/09489039A  
; Patent No. 6610836  
; GENERAL INFORMATION:  
; APPLICANT: Gary Breton et. al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA  
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 2709.2004001  
; CURRENT APPLICATION NUMBER: US/09/489,039A  
; CURRENT FILING DATE: 2000-01-27  
; PRIOR APPLICATION NUMBER: US 60/117,747  
; PRIOR FILING DATE: 1999-01-29  
; NUMBER OF SEQ ID NOS: 14342  
; SEQ ID NO 14144  
; LENGTH: 527  
; TYPE: PRT  
; ORGANISM: Klebsiella pneumoniae  
US-09-489-039A-14144

Query Match 54.5%; Score 36; DB 4; Length 527;  
Best Local Similarity 100.0%; Pred. No. 1.4e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 TELHERL 13  
||| |||||  
Db 111 TELHERL 117

RESULT 41  
US-09-902-540-12939  
; Sequence 12939, Application US/09902540  
; Patent No. 6833447  
; GENERAL INFORMATION:  
; APPLICANT: Goldman, Barry S.  
; APPLICANT: Hinkle, Gregory J.  
; APPLICANT: Slater, Steven C.  
; APPLICANT: Wiegand, Roger C.  
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof  
; FILE REFERENCE: 38-10(15849)B  
; CURRENT APPLICATION NUMBER: US/09/902,540  
; CURRENT FILING DATE: 2001-07-10  
; PRIOR APPLICATION NUMBER: 60/217,883  
; PRIOR FILING DATE: 2000-07-10

```
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 12939
; LENGTH: 715
; TYPE: PRT
; ORGANISM: Myxococcus xanthus
US-09-902-540-12939

Query Match          54.5%; Score 36; DB 4; Length 715;
Best Local Similarity 63.6%; Pred. No. 2e+02;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy      2 LVPFATELHER 12
Db      570 LVPFAGSRHDR 580

RESULT 42
US-08-510-646B-33
; Sequence 33, Application US/08510646B
; Patent No. 6077699
; GENERAL INFORMATION:
; APPLICANT: Blanc, Veronique
; APPLICANT: Blanc, Francis
; APPLICANT: Crouzet, Joel
; APPLICANT: Jacques, Nathalie
; APPLICANT: Lacroix, Patricia
; APPLICANT: Thibaut, Denis
; APPLICANT: Zagorec, Monique
; APPLICANT: Debussche, Laurent
; APPLICANT: De Crecy-Lagard, Valerie
; TITLE OF INVENTION: Polypeptides Involved In The
; TITLE OF INVENTION: Biosynthesis Of Streptogramins, Nucleotide Sequences
; TITLE OF INVENTION: Coding For These Polypeptides And Their Use
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett & Dunner
; STREET: 1300 I Street, N.W., Suite 700
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/510,646B
; FILING DATE: 03-AUG-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/403,852
; FILING DATE: 10-MAY-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/FR 93/00923
; FILING DATE: 25-SEP-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 92/11441
; FILING DATE: 25-SEP-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Meyers, Kenneth J.
; REGISTRATION NUMBER: 25,146
; REFERENCE/DOCKET NUMBER: 03806.0054-01000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 408-4000
; TELEFAX: (202) 408-4400
; INFORMATION FOR SEQ ID NO: 33:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 739 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-510-646B-33
```

```
Query Match          54.5%; Score 36; DB 3; Length 739;
Best Local Similarity 60.0%; Pred. No. 2.1e+02;
Matches 9; Conservative 1; Mismatches 3; Indels 2; Gaps 1;

Qy      1 KLVPF--ATELHERL 13
Db      267 RTVPFHLPELHERL 281

RESULT 43
US-09-270-767-35522
; Sequence 35522, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 35522
; LENGTH: 141
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-35522

Query Match          53.0%; Score 35; DB 4; Length 141;
Best Local Similarity 62.5%; Pred. No. 52;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy      3 VPFATELH 10
Db      40 IPFSSELH 47

RESULT 44
US-09-270-767-50739
; Sequence 50739, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 50739
; LENGTH: 141
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-50739

Query Match          53.0%; Score 35; DB 4; Length 141;
Best Local Similarity 62.5%; Pred. No. 52;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy      3 VPFATELH 10
Db      40 IPFSSELH 47

RESULT 45
US-09-134-001C-2941
; Sequence 2941, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
```

```

; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 2941
; LENGTH: 143
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-2941

Query Match      53.0%; Score 35; DB 3; Length 143;
Best Local Similarity 58.3%; Pred. No. 52;
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY      1 KLVPFATLHER 12
Db      64 KLTFFLELEKR 75

```

Search completed: September 1, 2005, 16:42:17  
Job time : 45 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 1, 2005, 16:23:14 ; Search time 39 Seconds  
(without alignments)  
32.072 Million cell updates/sec

Title: US-09-993-366-1

Perfect score: 66

Sequence: 1 KLVPPATELHERL 13

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database :

PIR 79:\*\*

1: Piri:\*\*

2: Piri:\*\*

3: Piri:\*\*

4: Piri:\*\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	66	100.0	396	1 LPHUM4	apolipoprotein A-I
2	66	100.0	401	2 A47141	apolipoprotein A-I
3	66	100.0	429	2 S29565	apolipoprotein A-I
4	41	62.1	261	2 AG3116	conserved hypotet
5	41	62.1	265	2 F98170	hypothetical prote
6	40	60.6	266	1 LPRB1Z	apolipoprotein A-I
7	39	59.1	241	2 A24998	apolipoprotein A-I
8	39	59.1	353	2 S70629	Rad51 homolog mei-
9	39	59.1	697	2 E96752	hypothetical prote
10	39	59.1	2796	2 JC4743	fatty-acid synthas
11	39	59.1	3069	2 H70656	fatty acid synthas
12	39	59.1	3076	2 A87058	fatty acid synthas
13	38	57.6	317	2 G70313	lipopolysaccharide
14	38	57.6	391	1 LPRTRA	apolipoprotein A-I
15	38	57.6	405	2 C75567	adenylosuccinate s
16	38	57.6	428	2 D84153	adenylosuccinate s
17	38	57.6	432	2 G97653	adenylosuccinate s
18	38	57.6	432	2 AE2877	adenylosuccinate s
19	38	57.6	473	2 AD0747	alpha,alpha-trehal
20	38	57.6	543	2 F69501	acetylactate synth
21	38	57.6	589	2 E71251	probable oligoendo
22	38	57.6	847	2 T52436	RF28 protein - Ara
23	38	57.6	850	2 JC5047	ras GTPase-activat
24	38	57.6	983	1 AG2381	Glycine cleavage s
25	38	57.6	987	1 TQBS30	transposase - Baci
26	38	57.6	1456	2 T01397	LTR gag/pol polypr
27	38	57.6	1474	2 B85188	retrotransposon li
28	38	57.6	2088	2 E71436	hypothetical prote
29	37	56.1	222	2 T26500	peptidylprolyl iso

30	37	56.1	231	2 JQ0704	apolipoprotein A-I
31	37	56.1	259	2 A24700	apolipoprotein A-I
32	37	56.1	264	2 S31394	apolipoprotein A-I
33	37	56.1	265	2 A46018	apolipoprotein AI
34	37	56.1	265	2 T06672	apolipoprotein A-I
35	37	56.1	289	2 C96752	probable DNA-bindi
36	37	56.1	339	2 S35642	RAD51 protein homo
37	37	56.1	339	2 S58295	RAD51 protein homo
38	37	56.1	339	2 A48221	RAD51 protein homo
39	37	56.1	394	2 S45881	probable purine nu
40	37	56.1	448	2 C83347	probable dipeptida
41	37	56.1	463	2 T36621	peptidyl prolyl ci
42	37	56.1	621	2 S37664	peptidomeric polypro
43	37	56.1	624	1 JC4510	pullulanase (EC 3.
44	37	56.1	630	2 S37663	peptidomeric polypro
45	37	56.1	1046	2 S39630	multidrug-efflux t
46	37	56.1	1046	2 E83593	RND multidrug effl
47	37	56.1	1154	1 VGIHIB	E2 glycoprotein pr
48	37	56.1	1162	1 VGIHAK	E2 glycoprotein pr
49	37	56.1	1162	2 S14940	E2 glycoprotein pr
50	37	56.1	1162	2 S14939	E2 glycoprotein pr
51	37	56.1	1162	2 S07421	E2 glycoprotein pr
52	37	56.1	1253	2 T42376	5-methyltetrahydro
53	37	56.1	1327	2 B06674	Aida-1 adhesin-lik
54	37	56.1	1349	2 E85524	probable beta-barr
55	36	54.5	162	2 S78127	NADH2 dehydrogenas
56	36	54.5	222	2 C86302	probable NADH-ubiq
57	36	54.5	222	2 S52380	NADH2 dehydrogenas
58	36	54.5	229	2 S52386	NADH2 dehydrogenas
59	36	54.5	229	2 S52385	NADH2 dehydrogenas
60	36	54.5	264	2 S60367	transcription fact
61	36	54.5	314	2 F83945	tRNA isopentenylpy
62	36	54.5	335	1 DEBSGF	glycerolaldehyde-3-p
63	36	54.5	419	2 AH2403	adenylosuccinate s
64	36	54.5	454	2 T01400	translation elonga
65	36	54.5	464	2 AF0214	probable membrane
66	36	54.5	518	2 A85793	probable transport
67	36	54.5	518	2 E90944	probable transport
68	36	54.5	518	2 A81865	conserved hypotet
69	36	54.5	518	2 C81077	conserved hypotet
70	36	54.5	518	2 H64942	hypothetical prote
71	36	54.5	519	2 AE0726	probable membrane
72	36	54.5	764	2 A72394	alpha-xylosidase -
73	36	54.5	879	2 AC2347	hypothetical prote
74	36	54.5	935	2 AE2469	two-component hybr
75	36	54.5	1053	1 GNLJBT	HIV-1 retropepsin
76	36	54.5	1057	1 OYMSAR	atrial natriuretic
77	36	54.5	1057	1 OYRTR	atrial natriuretic
78	36	54.5	1057	2 I55319	guanylyl cyclase A
79	36	54.5	1057	2 I57963	natriuretic peptid
80	36	54.5	1382	2 T01789	protoporphylin IX
81	36	54.5	1987	2 T30874	virginiamycin S sy
82	36	54.5	7463	2 T36248	CNA peptidase synthe
83	35.5	53.8	3344	2 JQ1899	genome polyprotein
84	35	53.0	89	2 I40752	abortive infection
85	35	53.0	164	2 S21830	apolipoprotein A-I
86	35	53.0	180	2 AC0056	conserved hypotet
87	35	53.0	286	2 C83474	hypothetical prote
88	35	53.0	284	2 JC5456	apolipoprotein A-I
89	35	53.0	266	1 LPDGA1	apolipoprotein A-I
90	35	53.0	279	2 A11399	ABC transporter (A
91	35	53.0	279	2 AH1775	ABC transporter (A
92	35	53.0	336	2 F90435	hypothetical prote
93	35	53.0	363	1 G69869	Xaa-Pro dipeptidase
94	35	53.0	394	2 G72212	exodeoxyribonuclea
95	35	53.0	401	2 T35956	probable acyl-CoA
96	35	53.0	474	2 A85803	trehalose-6-phosph
97	35	53.0	474	2 D90954	trehalose-6-phosph
98	35	53.0	474	2 D90954	alpha, alpha-trehal
99	35	53.0	492	2 T09754	catalase (EC 1.11.
100	35	53.0	521	1 B35441	alkyl hydroperoxid

## ALIGNMENTS

```

RESULT 1
LPHU44
apolipoprotein A-IV precursor [validated] - human
N/Alternate names: apoA-IV
C/Species: Homo sapiens (man)
C/Date: 30-Sep-1987 #sequence revision 30-Sep-1987 #text change 09-Jul-2004
C/Accession: A94137; A94059; A29330; A26280; I37177; C54223; A61203; A26481; S02
R;Karathanasis, S.K.; Oetgen, P.; Haddad, I.A.; Antonarakis, S.E.
Proc. Natl. Acad. Sci. U.S.A. 83, 8457-8461, 1986
A/Title: Structure, evolution, and polymorphisms of the human apolipoprotein A4 gene (AF
A/Reference number: A94137; MUID:87041474; PMID:3095836
A/Accession: A94137
A/Molecule type: DNA
A/Residues: 1-396 <KAR1>
A/Cross-references: UNIPROT:P06727; UNIPROT:Q13784; GB:M14642; NID:gl78760; PIDN:AAA5174
R;Karathanasis, S.K.
Proc. Natl. Acad. Sci. U.S.A. 82, 6374-6378, 1985
A/Title: Apolipoprotein multigene family: tandem organization of human apolipoprotein A
A/Reference number: A94059; MUID:86016704; PMID:3931073
A/Accession: A94059
A/Molecule type: mRNA
A/Residues: 135-378 <KAR2>
A/Cross-references: GB:M10373; NID:G563319; PIDN:AAB59516.1; PID:G563320
R;Karathanasis, S.K.; Yunis, I.
Biochemistry 25, 3962-3970, 1986
A/Title: Structure, evolution, and tissue-specific synthesis of human apolipoprotein AIV
A/Reference number: A24449; MUID:86296629; PMID:3755616
A/Accession: A24449
A/Molecule type: mRNA
A/Residues: 1-396 <KAR3>
A/Cross-references: GB:M13654; NID:gl78758; PIDN:AAA51744.1; PID:gl78759
R;Gordon, J.I.; Bisgaier, C.L.; Sims, H.F.; Sachdev, O.P.; Glickman, R.M.; Strauss, A.W.
J. Biol. Chem. 259, 468-474, 1984
A/Title: Biosynthesis of human preapolipoprotein A-IV.
A/Reference number: A92475; MUID:84161950; PMID:6706947
A/Contents: annotation; signal sequence cleavage site
R;Elshourbagy, N.A.; Walker, D.W.; Paik, Y.K.; Boguski, M.S.; Freeman, M.; Gordon, J.I.;
J. Biol. Chem. 262, 7973-7981, 1987
A/Title: Structure and expression of the human apolipoprotein A-IV gene.
A/Reference number: A29330; MUID:87250378; PMID:3036793
A/Accession: A29330
A/Molecule type: DNA
A/Residues: 1-157, 'T', 159, 'Y', 161-278, 'R', 280-326, 'T', 328-379, 'H', 381-396 <EL>
A/Cross-references: GB:J02758; NID:gl78756; PIDN:AAA96733.1; PID:gl78757
R;Elshourbagy, N.A.; Walker, D.W.; Boguski, M.S.; Gordon, J.I.; Taylor, J.M.
J. Biol. Chem. 261, 1998-2002, 1986
A/Title: The nucleotide and derived amino acid sequence of human apolipoprotein A-IV mRNA
A/Reference number: A26280; MUID:86111885; PMID:3080432
A/Accession: A26280
A/Molecule type: mRNA
A/Residues: 21-157, 'T', 159, 'Y', 161-278, 'R', 280-326, 'T', 328-379, 'H', 381-396 <EL2>
A/Cross-references: GB:M14566; NID:gl78778; PIDN:AAA51748.1; PID:gl78779
R;Yang, C.Y.; Gu, Z.W.; Chong, I.S.; Xiong, W.J.; Rosseneu, M.; Yang, H.X.; Lee, B.R.; G
Biochim. Biophys. Acta 1002, 231-237, 1989
A/Title: The primary structure of human apolipoprotein A-IV.
A/Reference number: I37177; MUID:89194198; PMID:2930771
A/Accession: I37177
A/Status: translated from GB/EMBL/DDBJ
A/Molecule type: mRNA
A/Residues: 1-157, 'T', 159, 'Y', 161-278, 'R', 280-396 <YANI>
A/Cross-references: EMBL:X13629; NID:g28761; PIDN:CAA31955.1; PID:g28762
A/Note: submitted to the EMBL Data Library, January 1989
R;Kunitake, S.T.; Carilli, C.T.; Lau, K.; Protter, A.A.; Naya-Vigne, J.; Kane, J.P.
Biochemistry 33, 1988-1993, 1994
A/Title: Identification of proteins associated with apolipoprotein A-I-containing lipop
A/Reference number: A54223; MUID:94162201; PMID:8117655
A/Accession: C54223
A/Molecule type: protein
A/Residues: 'X', 22, 'X', 24, 'X', 26-31, 'X', 33-34 <KUN>
R;Tenkanen, H.; Lukka, M.; Jauhainen, M.; Metso, J.; Baumann, M.; Peltonen, L.; Ehnholm

```

```

Arterioscler. Thromb. 11, 851-856, 1991
A/Title: The mutation causing the common apolipoprotein A-IV polymorphism is a glutamin
A/Reference number: A61203; MUID:91291788; PMID:2065039
A/Accession: A61203
A/Status: not compared with conceptual translation
A/Molecule type: DNA
A/Residues: 378-379, 'H', 381-382 <TEN>
C/Comment: ApoA-IV is a major lipoprotein of lymph chylomicrons. In human plasma, it is
C/Note: ApoA-IV is synthesized primarily in the intestine.
C/Genetics:
A/Gene: GDB:AP044
A/Cross-references: GDB:119000; OMIM:107690
A/Map position: 11q23-11q23
A/Introns: 17/1; 59/2
C/Superfamily: apolipoprotein A-I
C/Keywords: chylomicron; HDL; intestine; lipid binding; lipid transport; lipoprotein; p
F;1-20/Domain: signal sequence #status experimental <SIG>
F;21-396/Product: apolipoprotein A-IV #status experimental <MAT>
Query Match 100.0%; Score 66; DB 1; Length 396;
Best Local Similarity 100.0%; Pred. No. 0.00041;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KLVPPFATELHERL 13
DB 79 KLVPPFATELHERL 91
RESULT 2
A47141
apolipoprotein A-IV I isoform - baboon (fragment)
C/Species: Papio sp. (baboon)
C/Date: 16-Feb-1994 #sequence revision 18-Nov-1994 #text change 12-Apr-1995
R;Hixson, J.E.; Kammerer, C.M.; Mott, G.E.; Britten, M.L.; Birnbaum, S.; Powers, P.K.;
J. Biol. Chem. 268, 15667-15673, 1993
A/Title: Baboon apolipoprotein A-IV. Identification of Lys76-->Glu that distinguishes t
A/Reference number: A47141; MUID:93340170; PMID:8101842
A/Accession: A47141
A/Status: preliminary
A/Molecule type: nucleic acid
A/Residues: 1-401 <HX>
A/Experimental source: intestine
A/Note: sequence extracted from NCBI backbone (NCBI:136009, NCBIP:136010)
C/Superfamily: apolipoprotein A-I
Query Match 100.0%; Score 66; DB 2; Length 401;
Best Local Similarity 100.0%; Pred. No. 0.00041;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KLVPPFATELHERL 13
DB 63 KLVPPFATELHERL 75
RESULT 3
S29565
apolipoprotein A-IV - crab-eating macaque
C/Species: Macaca fascicularis (crab-eating macaque)
C/Date: 06-Jan-1995 #sequence revision 06-Jan-1995 #text change 09-Jul-2004
R;Osada, J.; Pocovi, M.; Nicolosi, R.J.; Schaefer, E.J.; Ordovas, J.M.
Biochim. Biophys. Acta 1172, 335-339, 1993
A/Title: Nucleotide sequences of the Macaca fascicularis apolipoprotein C-III and A-IV
A/Reference number: S30195; MUID:93192330; PMID:8448212
A/Accession: S30195
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-429 <OSA>
A/Cross-references: UNIPROT:P33621; EMBL:X68361; NID:g38050; PIDN:CAA48421.1; PID:g3805
C/Genetics:
A/Introns: 17/1; 59/2
C/Superfamily: apolipoprotein A-I

```





A; Introns: 32/3; 82/1

C; Superfamily: yeast DNA repair protein RAD51

Query Match 59.1%; Score 39; DB 2; Length 353;

Best Local Similarity 56.2%; Pred. No. 27;

Matches 9; Conservative 2; Mismatches 1; Indels 4; Gaps 1;

QY 1 KLVPF---ATELHER 12

|||||

84 KLVPMGFTTATEMHQR 99

RESULT 9

E96752

hypothetical protein F28P22.5 [imported] - Arabidopsis thaliana

C; Species: Arabidopsis thaliana (mouse-ear cress)

C; Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 09-Jul-2004

C; Accession: E96752

R; Theologidis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, C.; Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.; Jensen, N.F.; Hughes, B.; Huizar, L.

Nature 408, 816-820, 2000

A; Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Mafti, R.; Marziani, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A; Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A; Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A; Reference number: A86141; MUID:21016719; PMID:11130712

A; Accession: E96752

A; Status: preliminary

A; Molecule type: DNA

A; Residues: 1-697 <STO>

A; Cross-references: UNIPROT:Q9CA19; GB:AE005173; NID:g6648152; PIDN:AAF21152.1; GSPDB:GN

C; Genetics:

A; Gene: F28P22.5

A; Map position: 1

Query Match

Best Local Similarity 59.1%; Score 39; DB 2; Length 697;

Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 KLVPFATELHERL 13

|||||

647 RLKEFATEDHERI 659

RESULT 10

JC4743

fatty-acid synthase (EC 2.3.1.85) - Mycobacterium bovis

C; Species: Mycobacterium bovis

C; Date: 10-May-1996 #sequence\_revision 16-Aug-1996 #text\_change 09-Jul-2004

C; Accession: JC4743

R; Fernandes, N.D.; Kolattukudy, P.E.

Gene 170, 95-99, 1996

A; Title: Cloning, sequencing and characterization of a fatty acid synthase-encoding gene

A; Reference number: JC4743; MUID:96200863; PMID:8621098

A; Accession: JC4743

A; Molecule type: DNA

A; Residues: 1-2796 <FER>

A; Cross-references: UNIPROT:Q48926; GB:U36763; NID:g1036834; PIDN:AA03809.1; PID:g10368

A; Note: the source is designated as Mycobacterium tuberculosis var. bovis BCG

C; Comment: This enzyme catalyzes both de novo synthesis and chain elongation of fatty ac

C; Genetics:

A; Gene: fas

C; Superfamily: Mycobacterium tuberculosis fatty-acid synthase

F; 2188-2193/Region: nucleotide binding #status predicted

F; 57/Active site: Ser #status predicted

F; 1693/Binding site: phosphopantetheine (Ser) (covalent) #status predicted

F; 2598/Active site: Cys #status predicted

Query Match

Best Local Similarity 59.1%; Score 39; DB 2; Length 2796;

Best Local Similarity 80.0%; Pred. No. 2.5e+02;

Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 PFATELHERL 13

|||||

Db 2755 PNETELHERL 2764

RESULT 11

H70656

fatty-acid synthase (EC 2.3.1.85) - Mycobacterium tuberculosis (strain H37RV)

C; Species: Mycobacterium tuberculosis

C; Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 09-Jul-2004

C; Accession: H70656

R; Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.T.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.

Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.

Nature 393, 537-544, 1998

A; Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.

A; Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome

A; Reference number: A70500; MUID:98295987; PMID:9634230

A; Accession: H70656

A; Status: preliminary; nucleic acid sequence not shown; translation not shown

A; Molecule type: DNA

A; Residues: 1-3069 <COL>

A; Cross-references: UNIPROT:P95029; GB:Z83863; GB:AL123456; NID:g3261685; PIDN:CAB06201

A; Experimental source: strain H37RV

C; Genetics:

A; Gene: fas

C; Superfamily: Brevibacterium ammoniagenes fatty-acid synthase

C; Keywords: acyltransferase; coenzyme A

Query Match

Best Local Similarity 59.1%; Score 39; DB 2; Length 3069;

Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 PFATELHERL 13

|||||

Db 2881 PNETELHERL 2890

RESULT 12

A87058

fatty acid synthase [imported] - Mycobacterium leprae

C; Species: Mycobacterium leprae

C; Date: 20-Apr-2001 #sequence\_revision 20-Apr-2001 #text\_change 09-Jul-2004

C; Accession: A87058

R; Cole, S.T.; Eigimeier, K.; Parkhill, J.; James, K.D.; Thomson, N.R.; Wheeler, P.R.; Holroyd, R.; Davies, R.M.; Devlin, K.; Duthoy, S.; Feltwell, T.; Fraser, A.; Hamlin, N.; Holroyd, S.

Nature 409, 1007-1011, 2001

A; Authors: Rutter, S.; Seeger, K.; Simon, S.; Simmonds, M.; Skelton, J.; Squares, R.; S.

A; Title: Massive gene decay in the leprosy bacillus.

A; Reference number: A86909; MUID:21128732; PMID:11234002

A; Accession: A87058

A; Status: preliminary

A; Molecule type: DNA

A; Residues: 1-3076 <STO>

A; Cross-references: UNIPROT:Q9X7E2; GB:AL450380; NID:g13093156; PIDN:CAC31572.1; GSPDB:G

C; Genetics:

C; Superfamily: Brevibacterium ammoniagenes fatty-acid synthase

Query Match

Best Local Similarity 59.1%; Score 39; DB 2; Length 3076;

Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 PFATELHERL 13

|||||

Db 2888 PNETELHERL 2897

RESULT 13

G70313  
lipopolysaccharide heptosyltransferase (BC 2.4.99.-) II rfaC2 [similarity] - Aquifex aeolicus  
C;Species: Aquifex aeolicus  
C;Date: 08-May-1998 #sequence\_revision 08-May-1998 #text\_change 09-Jul-2004  
C;Accession: G70313  
R;Decker, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; O'Neil, J.  
Nature 392, 353-358, 1998  
A;Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.  
A;Reference number: A70300; MUID:98196666; PMID:9537320  
A;Accession: G70313  
A;Status: preliminary; nucleic acid sequence not shown; translation not shown  
A;Molecule type: DNA  
A;Residues: 1-317 <AQF>  
A;Cross-references: UNIPROT:O66538; GB:AE000675; NID:g2982863; PIDN:AAC06484.1; PID:g2982863  
A;Experimental source: strain VF5  
C;Genetics:  
A;Gene: rfaC2  
C;Superfamily: ADP-heptose-LPS heptosyltransferase II  
C;Keywords: glycosyltransferase  
Query Match 57.6%; Score 38; DB 2; Length 317;  
Best Local Similarity 72.7%; Pred. No. 37;  
Matches 8; Conservative 0; Mismatches 0; Indels 3; Gaps 0;  
QY 1 KLVPFATELHE 11  
|||||  
Db 110 KLVPHRWELHE 120  
RESULT 14  
LPRTA4  
apolipoprotein A-IV precursor - rat  
C;Species: Rattus norvegicus (Norway rat)  
C;Date: 27-Nov-1985 #sequence\_revision 27-Nov-1985 #text\_change 09-Jul-2004  
C;Accession: A03095; A25214; C24700  
R;Boguski, M.S.; Elshourbagy, N.; Taylor, J.M.; Gordon, J.I.  
Proc. Natl. Acad. Sci. U.S.A. 81, 5021-5025, 1984  
A;Title: Rat apolipoprotein A-IV contains 13 tandem repetitions of a 22-amino acid segment.  
A;Reference number: A03095; MUID:84298074; PMID:6591177  
A;Accession: A03095  
A;Molecule type: mRNA  
A;Residues: 1-391 <BOG>  
A;Cross-references: UNIPROT:P02651; GB:M00002; GB:K02421; NID:g202949; PIDN:AAA85909.1;  
R;Boguski, M.S.; Birkenmeier, E.H.; Elshourbagy, N.A.; Taylor, J.M.; Gordon, J.I.  
J. Biol. Chem. 261, 6398-6407, 1986  
A;Title: Evolution of the apolipoproteins. Structure of the rat APO-A-IV gene and its relationship to the human APO-A-IV gene.  
A;Reference number: A25214; MUID:86196059; PMID:3009456  
A;Accession: A25214  
A;Molecule type: protein  
A;Residues: 1-252, 'Q', 254-391 <BO2>  
R;Haddad, I.A.; Ordovas, J.M.; Fitzpatrick, T.; Karathanasis, S.K.  
J. Biol. Chem. 261, 13268-13277, 1986  
A;Title: Linkage, evolution, and expression of the rat apolipoprotein A-I, C-III, and A-IV genes.  
A;Reference number: A92558; MUID:87008540; PMID:3020028  
A;Accession: C24700  
A;Molecule type: DNA  
A;Residues: 1-252, 'Q', 254-391 <HAD>  
A;Cross-references: GB:J02588; NID:g202937; PIDN:AAA40747.1; PID:g202941  
C;Comment: This apoprotein is a major component of HDL and chylomicrons but, unlike other apolipoproteins, it is not a lipid carrier.  
C;Comment: Nine of the thirteen 22-amino acid tandem repeats (each 22-mer is actually a 23-mer) are amphipathic. They may therefore serve as lipid-binding sites.  
C;Superfamily: apolipoprotein A-I  
C;Keywords: chylomicron; duplication; HDL; lipid transport; plasma; tandem repeat  
F;1-20/Domain: signal sequence #status predicted <SIG>  
F;21-391/Product: apolipoprotein A-IV #status predicted <MAT>  
F;33-330/Region: 22-residue repeats  
Query Match 57.6%; Score 38; DB 1; Length 391;  
Best Local Similarity 63.6%; Pred. No. 46;  
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
QY 2 LVPPATELHER 12  
|||||

Db 179 MVPFANELKEK 189  
|||||  
RESULT 15  
C75567  
adenylosuccinate synthase - Deinococcus radiodurans (strain R1)  
C;Species: Deinococcus radiodurans  
C;Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 09-Jul-2004  
C;Accession: C75567  
R;White, O.; Eisen, J.A.; Heideberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.; M.; Sheth, H.O.; Venter, J.C.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; M. S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.  
Science 286, 1571-1577, 1999  
A;Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.  
A;Reference number: A75250; MUID:20036896; PMID:10567266  
A;Accession: C75567  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-405 <WHI>  
A;Cross-references: UNIPROT:Q9RYB5; GB:AE001867; GB:AE000513; NID:g6457693; PIDN:AAFO96  
C;Genetics:  
A;Experimental source: strain R1  
A;Gene: DR0035  
A;Map position: 1  
C;Superfamily: adenylosuccinate synthase  
Query Match 57.6%; Score 38; DB 2; Length 405;  
Best Local Similarity 75.0%; Pred. No. 48;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
QY 4 PFATELHE 11  
|||||  
Db 266 PFATEVHD 273  
RESULT 16  
D84153  
adenylosuccinate synthetase purA [imported] - Bacillus halodurans (strain C-125)  
C;Species: Bacillus halodurans  
C;Date: 01-Dec-2000 #sequence\_revision 01-Dec-2000 #text\_change 09-Jul-2004  
C;Accession: D84153  
R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, N.; Fujii, F.; Hiraoka, Y.  
Nucleic Acids Res. 28, 4317-4331, 2000  
A;Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and its relationship to the Bacillus pasteurii genome.  
A;Reference number: A83650; MUID:20512582; PMID:11058132  
A;Accession: D84153  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-428 <STO>  
A;Cross-references: UNIPROT:Q9K5R0; GB:AP001520; GB:BA000004; NID:g10176401; PIDN:BAB07  
A;Experimental source: strain C-125  
C;Genetics:  
A;Gene: purA  
C;Superfamily: adenylosuccinate synthase  
Query Match 57.6%; Score 38; DB 2; Length 428;  
Best Local Similarity 60.0%; Pred. No. 51;  
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
QY 4 PFATELHERL 13  
|||||  
Db 276 PFPELHDEI 285  
RESULT 17  
G97653  
adenylosuccinate synthetase (IMP-aspartate ligase) (adss) [imported] - Agrobacterium tumefaciens  
C;Species: Agrobacterium tumefaciens  
C;Date: 30-Sep-2001 #sequence\_revision 30-Sep-2001 #text\_change 09-Jul-2004  
C;Accession: G97653  
R;Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Quorllo, B.; Goldmar, A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markeiz, B.

Science 294, 2323-2328, 2001  
A;Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent *Agrobacterium tumefaciens* strain C58  
A;Reference number: A97359; MUID:21608551; PMID:11743194  
A;Accession: G97653  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-432 <KUR>  
A;Cross-references: UNIPROT:Q8UCN6; GB:AE007869; PIDN:AAK88184.1; PID:g15157630; GSPDB:C58  
C;Genetics:  
A;Gene: AGR\_C\_4442  
A;Map position: circular chromosome  
C;Superfamily: adenylosuccinate synthase

Query Match 57.6%; Score 38; DB 2; Length 432;  
Best Local Similarity 60.0%; Pred. No. 51;  
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 4 PFATELHERL 13  
DB 279 PPFTLHDEI 288  
||| ||||| :  
||| ||||| :

RESULT 18  
AE2877  
adenylosuccinate synthetase [imported] - *Agrobacterium tumefaciens* (strain C58, Dupont)  
C;Species: *Agrobacterium tumefaciens*  
C;Date: 11-Jan-2002 #sequence\_revision 11-Jan-2002 #text\_change 09-Jul-2004  
C;Accession: AE2877  
R;Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I.; Erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutyavin, T.; Levy, R.; Li, M.; McClellan, J.; Karp, P.; Romero, P.; Zhang, S.  
Science 294, 2317-2323, 2001  
A;Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm, J.; E.W.  
A;Title: The Genome of the Natural Genetic Engineer *Agrobacterium tumefaciens* C58.  
A;Reference number: AB2577; MUID:21608550; PMID:11743193  
A;Accession: AE2877  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-432 <KUR>  
A;Cross-references: UNIPROT:Q8UCN6; GB:AE008688; PIDN:AAL43435.1; PID:g17740937; GSPDB:C58  
C;Genetics:  
A;Gene: pura  
A;Map position: circular chromosome  
C;Superfamily: adenylosuccinate synthase

Query Match 57.6%; Score 38; DB 2; Length 432;  
Best Local Similarity 60.0%; Pred. No. 51;  
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 4 PFATELHERL 13  
DB 279 PPFTLHDEI 288  
||| ||||| :  
||| ||||| :

RESULT 19  
AD0747  
alpha,alpha-trehalose-phosphate synthase (UDP-forming) (EC 2.4.1.15) - *Salmonella enterica* serovar Typhi  
C;Species: *Salmonella enterica* subsp. *enterica* serovar Typhi  
A;Note: this species has also been called *Salmonella typhi*  
C;Date: 09-Nov-2001 #sequence\_revision 09-Nov-2001 #text\_change 18-Nov-2002  
C;Accession: AD0747  
R;Farhnik, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar, J.; Moule, S.; O'Gaora, P.  
Nature 413, 848-852, 2001  
A;Authors: Parry, C.; Ouail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; A;Title: Complete genome sequence of a multiple drug resistant *Salmonella enterica* serovar Typhi  
A;Reference number: AB0502; MUID:21534947; PMID:11677608  
A;Accession: AD0747  
A;Status: preliminary  
A;Molecule type: DNA

A;Residues: 1-473 <PAR>  
A;Cross-references: GB:AL513382; PIDN:CAD05679.1; PID:g16503174; GSPDB:GN00176  
C;Genetics:  
A;Gene: STY2137  
C;Superfamily: Schizosaccharomyces pombe alpha,alpha-trehalose-phosphate synthase (UDP-forming)  
C;Keywords: glycosyltransferase; hexosyltransferase

Query Match 57.6%; Score 38; DB 2; Length 473;  
Best Local Similarity 63.6%; Pred. No. 57;  
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 LVFPATELHER 12  
DB 133 LLPFASLEKR 143  
||| ||||| :  
||| ||||| :

RESULT 20  
F69501  
acetylactate synthase, large subunit (ilvB-3) homolog - *Archaeoglobus fulgidus*  
C;Species: *Archaeoglobus fulgidus*  
C;Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 16-Aug-2004  
C;Accession: F69501  
R;Kienk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson, J.; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.; Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.; Nature 390, 364-370, 1997  
A;Authors: Uterback, T.; Cotton, M.D.; Spriggs, T.; Artiach, P.; Kaine, B.P.; Sykes, S.; Smith, H.O.; Woese, C.R.; Venter, J.C.  
A;Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaeon *Pyrococcus furiosus*  
A;Reference number: A69250; MUID:198049343; PMID:9389475  
A;Accession: F69501  
A;Status: preliminary; nucleic acid sequence not shown; translation not shown  
A;Molecule type: DNA  
A;Residues: 1-543 <KLE>  
A;Cross-references: UNIPROT:028264; GB:AE000963; GB:AE000782; MID:g26689286; PIDN:AAB892  
C;Superfamily: Acetylactate synthase, large subunit/pyruvate oxidase; thiamin pyrophosphatase

Query Match 57.6%; Score 38; DB 2; Length 543;  
Best Local Similarity 63.6%; Pred. No. 66;  
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 KLVFPATELHE 11  
DB 126 RLVPETHRLHE 136  
||| ||||| |||||  
||| ||||| |||||

RESULT 21  
E71251  
probable oligonucleotide hydrolase F - *Syphilis spirochete*  
C;Species: *Treponema pallidum* subsp. *pallidum* (syphilis spirochete)  
C;Date: 24-Jul-1998 #sequence\_revision 24-Jul-1998 #text\_change 09-Jul-2004  
C;Accession: E71251  
R;Fraser, C.M.; Norris, S.J.; Weinstock, G.M.; White, O.; Sutton, G.G.; Dodson, R.; Gwinn, L.; Khakhria, H.; Richardson, D.; Howell, J.K.; Chidambaram, M.; Uterback, T.; McDaniel, L.; Science 281, 375-388, 1998  
A;Title: Complete genome sequence of *Treponema pallidum*, the syphilis spirochete.  
A;Reference number: A71250; MUID:98332770; PMID:9665876  
A;Accession: E71251  
A;Status: preliminary; nucleic acid sequence not shown; translation not shown  
A;Molecule type: DNA  
A;Residues: 1-589 <COL>  
A;Cross-references: UNIPROT:083989; GB:AE001270; GB:AE000520; MID:g33233350; PIDN:AAC659  
A;Experimental source: strain Nichols  
C;Genetics:  
A;Gene: TP1026

Query Match 57.6%; Score 38; DB 2; Length 589;  
Best Local Similarity 70.0%; Pred. No. 72;  
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 VFPATELHER 12  
DB 111 ||| ||||| :  
||| ||||| :

Db 246 VPLAELHEQ 255

## RESULT 22

RF28 protein - Arabidopsis thaliana retrotransposon (fragment)  
T52436  
C;Species: Arabidopsis thaliana (mouse-ear cress)  
C;Date: 24-Oct-2000 #sequence\_revision 24-Oct-2000 #text\_change 09-Jul-2004  
C;Accession: T52436  
R;Kawahara, A.; Kato, A.; Komeda, Y.  
submitted to the Protein Sequence Database, October 2000  
A;Reference number: Z26076  
A;Accession: T52436  
A;Molecule type: mRNA  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Residues: 1-847 <KUN>  
A;Cross-references: UNIPROT:Q9SLUA; EMBL:AB028223; PIDN:BA87949.1  
A;Experimental source: cultivar Columbia  
C;Genetics:  
A;Map position: 1  
A;Note: RF28  
C;Superfamily: retrovirus-related polyprotein

Query Match 57.6%; Score 38; DB 2; Length 847;  
Best Local Similarity 70.0%; Pred. No. 1.1e+02;  
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 PFATELHERL 13

Db 159 PTLTEIHERL 168

## RESULT 23

JC5047  
ras GTPase-activating protein - human  
C;Species: Homo sapiens (man)  
C;Date: 31-Jan-1997 #sequence\_revision 31-Jan-1997 #text\_change 09-Jul-2004  
C;Accession: JC5047  
R;Kobayashi, M.; Masui, T.; Kusuda, J.; Kameoka, Y.; Hashimoto, K.; Iwashita, S.  
Gene 175, 173-177, 1996  
A;Title: Human rasGTPase-activating protein (human counterpart of GAP1m): Sequence of the  
A;Reference number: JC5047; MUID:97074668; PMID:8917095  
A;Accession: JC5047  
A;Molecule type: mRNA  
A;Residues: 1-850 <KOB>  
A;Cross-references: UNIPROT:Q15283; DBJ:D78155; NID:q1060908; PIDN:BA11230.1; PID:d101  
C;Comment: This protein plays a role in the regulation of cell growth and differentiation  
C;Genetics:  
A;Gene: GAP1m  
A;Map position: 3q24-26  
F;356-568/Domain: ras-specific GAP catalytic domain homology <GAP>  
F;603-704/Domain: pleckstrin repeat homology <PK>

Query Match 57.6%; Score 38; DB 2; Length 850;  
Best Local Similarity 72.7%; Pred. No. 1.1e+02;  
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 KLVPFATELHE 11

Db 373 KLVPFATAVAE 383

## RESULT 24

AG2381  
glycine cleavage system protein P [imported] - Nostoc sp. (strain PCC 7120)  
C;Species: Nostoc sp. PCC 7120  
A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120  
C;Date: 14-Dec-2001 #sequence\_revision 14-Dec-2001 #text\_change 09-Jul-2004  
C;Accession: AG2381  
R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriuchih  
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S  
DNA Res. 8, 205-213, 2001  
A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana

A;Reference number: AB1807; MUID:21595285; PMID:11759840

A;Accession: AG2381

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-983 <KUR>

A;Cross-references: UNIPROT:QBYNE9; GB:BA000019; PIDN:BA876306.1; PID:gl7133744; GSPDB:

A;Experimental source: strain PCC 7120

C;Genetics:

A;Gene: all4607

Query Match 57.6%; Score 38; DB 2; Length 983;  
Best Local Similarity 61.5%; Pred. No. 1.2e+02;  
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 KLVPFATELHERL 13

Db 64 KTVPSAIRLHEQL 76

## RESULT 25

TQBS30  
transposase - Bacillus thuringiensis transposon Tn4430  
C;Species: Bacillus thuringiensis  
C;Date: 31-Mar-1992 #sequence\_revision 31-Mar-1992 #text\_change 09-Jul-2004  
C;Accession: S00570; S02241  
R;Mahillon, J.; Lereclus, D.  
EMBO J. 7, 1515-1526, 1988  
A;Title: Structural and functional analysis of Tn4430: identification of an integrase-1  
A;Reference number: S00554; MUID:88312602; PMID:2842151  
A;Accession: S00570  
A;Molecule type: DNA  
A;Residues: 1-987 <MAHL>  
A;Cross-references: UNIPROT:P10021; EMBL:X07651; NID:g40347; PIDN:CAA30492.1; PID:g4034

R;Mahillon, J.; Seurinck, J.

Nucleic Acids Res. 16, 11827-11828, 1988

A;Title: Complete nucleotide sequence of pGI2, a Bacillus thuringiensis plasmid contain

A;Reference number: S02047; MUID:89098342; PMID:3211758

A;Accession: S02241

A;Status: translation not shown

A;Molecule type: DNA

A;Residues: 1-987 <MAH2>

A;Cross-references: EMBL:X13481; NID:g3171732; PIDN:CAA31833.1; PID:g40318

C;Genetics:

A;Gene: tnpA

C;Superfamily: transposase Tn21

C;Keywords: DNA binding; DNA replication

Query Match 57.6%; Score 38; DB 1; Length 987;  
Best Local Similarity 50.0%; Pred. No. 1.2e+02;  
Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 2 LVFPFATELHERL 13

Db 288 LIDYATIEIHDRL 299

## RESULT 26

TQ1397  
LTR gag/pol polyprotein homolog T419.16 - Arabidopsis thaliana  
C;Species: Arabidopsis thaliana (mouse-ear cress)  
C;Date: 12-Feb-1999 #sequence\_revision 12-Feb-1999 #text\_change 09-Jul-2004  
C;Accession: TQ1397  
R;Parnell, L.D.; Gnoj, L.; de la Bastide, M.; Hameed, A.; Habermann, K.; Schutz, K.; Hu  
submitted to the EMBL Data Library, May 1998  
A;Description: Genomic sequence of BAC T419 from Arabidopsis thaliana, chromosome IV, n  
A;Reference number: Z14314  
A;Accession: TQ1397  
A;Status: translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-1456 <PAR>

A;Cross-references: UNIPROT:Q9T94; EMBL:AF069442; NID:g3242970; PID:g3924609

A;Experimental source: cultivar Columbia

C;Genetics:

A:Map position: 4  
A:Introns: 129/1  
A:Note: T419.16  
C:Superfamily: retrovirus-related polyprotein

Query Match 57.6%; Score 38; DB 2; Length 1456;  
Best Local Similarity 70.0%; Pred. No. 1.9e+02;  
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 PFATELHERL 13  
| | | | |  
DB 175 PSITEIHERL 184

RESULT 27  
B85188  
retrotransposon like protein [imported] - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cross)  
C:Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 09-Jul-2004  
C:Accession: B85188  
R:anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Spring Nature 402, 769-777, 1999  
A:Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.  
A:Reference number: A85001; MUID:20083488; PMID:10617198  
A:Accession: B85188  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-1474 <STO>  
A:Cross-references: UNIPROT:O23529; GB:NC\_001268; NID:G5302802; PIDN:CAB46043.1; GSPDB:G5302802; CESP:G5302802  
C:Genetics:  
A:Map position: 4  
A:Gene: dl4465c  
C:Superfamily: retrovirus-related polyprotein

Query Match 57.6%; Score 38; DB 2; Length 1474;  
Best Local Similarity 70.0%; Pred. No. 1.9e+02;  
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 PFATELHERL 13  
| | | | |  
DB 202 PSITEIHERL 211

RESULT 28  
E71436  
hypothetical protein - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cross)  
A:Variety: columbia  
C:Date: 03-Aug-1998 #sequence\_revision 03-Aug-1998 #text\_change 09-Jul-2004  
C:Accession: E71436  
R:Bevan, M.; Bancroft, I.; Bent, E.; Love, K.; Goodman, H.; Dean, C.; Bergkamp, R.; Dirk P.; Wedler, H.; Wedler, E.; Wambutt, R.; Weitzenger, T.; Pohl, T.M.; Terry, N.; Gielavanagh, T.; Hempel, S.; Kotter, P.; Entian, K.D.; Rieger, M.; Schaeffer, M.; Funk, B. Nature 391, 485-488, 1998  
A:Authors: Mueller-Auer, S.; Silvey, M.; James, R.; Montfort, A.; Pons, A.; Puigdomenech erhoft, A.; Moores, T.; Jones, J.D.G.; Eneva, T.; Palme, K.; Benes, V.; Rechman, S.; Andre Chwalatzi, N.  
A:Title: Analysis of 1.9 Mb of contiguous sequence from chromosome 4 of Arabidopsis thaliana  
A:Reference number: A71400; MUID:98121113; PMID:9461215  
A:Accession: E71436  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-2088 <BEV>  
A:Cross-references: UNIPROT:O23529; GB:297342; NID:G2245031; PID:G2245044  
C:Genetics:  
A:Map position: 4COP9-4G3845

Query Match 57.6%; Score 38; DB 2; Length 2088;  
Best Local Similarity 70.0%; Pred. No. 2.8e+02;  
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 PFATELHERL 13  
| | | | |

DB 886 PSITEIHERL 895

RESULT 29  
T26500  
peptidylprolyl isomerase (EC 5.2.1.8) Y17G7B.9 [similarity] - Caenorhabditis elegans  
N:Contains: cyclophilin  
C:Species: Caenorhabditis elegans  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
C:Accession: T26500  
R:Smye, R.  
submitted to the EMBL Data Library, June 1998  
A:Reference number: Z20225  
A:Accession: T26500  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-222 <WIL>  
A:Cross-references: UNIPROT:Q9XX17; EMBL:AL023828; PIDN:CAAL19454.1; GSPDB:GN000020; CESP:GN000020  
A:Experimental source: clone Y17G7B  
C:Genetics:  
A:Gene: CESP:Y17G7B.9  
A:Map position: 2  
A:Introns: 14/3; 60/1; 84/3  
C:Superfamily: peptidylprolyl isomerase; cyclophilin homology  
C:Keywords: cis-trans-isomerase

Query Match 56.1%; Score 37; DB 2; Length 222;  
Best Local Similarity 60.0%; Pred. No. 38;  
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 4 PFATELHERL 13  
| | | | |  
DB 82 PFKDEIHOQL 91

RESULT 30  
JQ0704  
apolipoprotein A-I - pig (fragment)  
C:Species: Sus scrofa domestica (domestic pig)  
C:Date: 31-Mar-1992 #sequence\_revision 31-Mar-1992 #text\_change 28-Oct-1994  
C:Accession: JQ0704  
R:Weiler-Guettler, H.; Sommerfeldt, M.; Papandriopoulou, A.; Mischek, U.; Bonitz, D.; J. Neurochem. 54, 444-450, 1990  
A:Title: Synthesis of apolipoprotein A-I in pig brain microvascular endothelial cells.  
A:Reference number: JQ0704; MUID:90132667; PMID:2105375  
A:Accession: JQ0704  
A:Molecule type: mRNA  
A:Residues: 1-231 <WEI>  
A:Note: the authors translated the codon CAG for residue 124 as His and GAC for residue 125  
C:Superfamily: apolipoprotein A-I  
C:Keywords: cholesterol metabolism; HDL; lipid binding; lipid transport; lipoprotein; p

Query Match 56.1%; Score 37; DB 2; Length 231;  
Best Local Similarity 61.5%; Pred. No. 40;  
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 KLVPPFATELHERL 13  
| | | | |  
DB 129 KLSPLAEELRDRL 141

RESULT 31  
A24700  
apolipoprotein A-I precursor - rat  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 30-Jun-1988 #sequence\_revision 30-Jun-1988 #text\_change 09-Jul-2004  
C:Accession: A24700; S00298; A05314  
R:Haddad, I.A.; Ordovas, J.M.; Fitzpatrick, T.; Karathanasis, S.K. J. Biol. Chem. 261, 13268-13277, 1986  
A:Title: Linkage, evolution, and expression of the rat apolipoprotein A-I, C-III, and A  
A:Reference number: A24700; MUID:87008540; PMID:3020028  
A:Accession: A24700  
A:Molecule type: DNA

A;Residues: 1-259 <HAD>  
A;Cross-references: UNIPROT:P04639; EMBL:J02597; NID:g202935; PIDN:AAA40745.1; PID:g20294  
R;Poncin, J.E.; Martial, J.A.; Gielen, J.E.  
Eur. J. Biochem. 140, 493-498, 1984  
A;Title: Cloning and structure analysis of the rat apolipoprotein A-I cDNA.  
A;Reference number: S00298; MUID:84207987; PMID:6426956  
A;Accession: S00298  
A;Molecule type: mRNA  
A;Residues: 1-259 <PON>  
A;Cross-references: GB:M00001; EMBL:X00558; NID:g202944; PIDN:AAA40749.1; PID:g202945  
R;Gordon, J.I.; Smith, D.P.; Andy, R.; Alpers, D.H.; Schonfeld, G.; Straus, A.W.  
J. Biol. Chem. 257, 971-978, 1982  
A;Title: The primary translation product of rat intestinal apolipoprotein A-I mRNA is an  
A;Reference number: A05314; MUID:82098162; PMID:6798036  
A;Accession: A05314  
A;Molecule type: protein  
A;Residues: 1-18, 'X', 20-21, 'X', 23-28, 'X', 30-38, 'X', 40-41, 'X', 43-45 <GOR>  
C;Comment: This protein is synthesized in the liver and small intestine. The propeptide  
C;Comment: This protein is a major component of the high density lipoproteins in plasma.  
C;Genetics:  
A;Introns: 15/1; 66/2  
C;Superfamily: apolipoprotein A-I  
C;Keywords: cholesterol metabolism; HDL; intestine; lipid binding; lipid transport; lipop  
F;1-18/Domain: signal sequence #status experimental <SIG>  
F;19-24/Domain: propeptide #status experimental <PRO>  
F;25-259/Product: apolipoprotein A-I #status experimental <MAT>  
Query Match 56.1%; Score 37; DB 2; Length 259;  
Best Local Similarity 63.6%; Pred. No. 45;  
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
QY 1 KLVPFATLH 11  
Db 141 KLSPLGLTGLHK 151  
RESULT 32  
S31394  
apolipoprotein A-I - pig  
C;Species: Sus scrofa domestica (domestic pig)  
C;Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 09-Jul-2004  
C;Accession: S31394  
R;Moehel, B.; Flach, R.; Weiss, B.; Weiler-Guetli, H.; Frey, A.; Zinke, H.; Gassen, H  
submitted to the EMBL Data Library, November 1992  
A;Description: Genomic organization of the porcine apolipoprotein AI gene and study of  
A;Reference number: S31394  
A;Accession: S31394  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-264 <MOE>  
A;Cross-references: UNIPROT:P18648; EMBL:X69477; NID:g1889; PIDN:CAA49234.1; PID:g1890  
C;Superfamily: apolipoprotein A-I  
Query Match 56.1%; Score 37; DB 2; Length 264;  
Best Local Similarity 61.5%; Pred. No. 46;  
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
QY 1 KLVPFATLH 13  
Db 162 KLSPLAEELRDL 174  
RESULT 33  
A46018  
apolipoprotein AI - pig  
C;Species: Sus scrofa domestica (domestic pig)  
C;Date: 21-Sep-1993 #sequence\_revision 18-Nov-1994 #text\_change 09-Jul-2004  
C;Accession: A46018  
R;Birchbauer, A.; Knipping, G.; Juritsch, B.; Aschauer, H.; Zechner, R.  
Genomics 15, 643-652, 1993  
A;Title: Characterization of the apolipoprotein AI and CIII genes in the domestic pig.  
A;Reference number: A46018; MUID:93224154; PMID:8468059  
A;Accession: A46018

A;Status: preliminary  
A;Molecule type: DNA; protein  
A;Residues: 1-265 <BIR>  
A;Cross-references: UNIPROT:P18648; GB:L00626; NID:g164358; PIDN:AAA30992.1; PID:g16435  
C;Note: sequence extracted from NCBI backbone (NCBI:129509, NCBI:P:129511)  
C;Superfamily: apolipoprotein A-I  
Query Match 56.1%; Score 37; DB 2; Length 265;  
Best Local Similarity 61.5%; Pred. No. 46;  
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;  
QY 1 KLVPFATLH 13  
Db 163 KLSPLAEELRDL 175  
RESULT 34  
JT0672  
apolipoprotein A-I - pig  
N;Alternate names: apo-A-I  
C;Species: Sus scrofa domestica (domestic pig)  
C;Date: 28-Oct-1994 #sequence\_revision 28-Oct-1994 #text\_change 09-Jul-2004  
C;Accession: JT0672; PN0471; A05311  
R;Frieu, V.N.; Patel, B.; Zhan, R.; Black, D.D.  
Gene 134, 267-270, 1993  
A;Title: Sequence of the porcine apoA-I gene.  
A;Reference number: JT0672; MUID:94085789; PMID:7916724  
A;Accession: JT0672  
A;Molecule type: DNA  
A;Residues: 1-265 <TRI>  
A;Cross-references: UNIPROT:P18648; EMBL:Z14124; NID:g1893  
A;Note: this translation is not annotated in GenBank entry SSAPOAIG, release 111.0; the  
R;Frieu, V.N.; Hasler-Rapacz, J.; Rapacz, J.; Black, D.D.  
Gene 123, 173-179, 1993  
A;Title: Sequences and expression of the porcine apolipoprotein A-I and C-III mRNAs.  
A;Reference number: PN0471; MUID:93154581; PMID:8428656  
A;Accession: PN0471  
A;Molecule type: mRNA  
A;Residues: 105-265 <TR2>  
A;Experimental source: liver  
R;Mahley, R.W.; Weisgraber, K.H.; Innerarity, T.; Brewer Jr., H.B.  
Biochemistry 15, 1928-1933, 1976  
A;Title: Characterization of the plasma lipoproteins and apoproteins of the Erythrocebu  
A;Reference number: A90395; MUID:76184721; PMID:178359  
A;Accession: A05311  
A;Molecule type: protein  
A;Residues: 25-34 <MAH>  
C;Comment: This protein is the major apolipoprotein of high-density lipoprotein and sei  
C;Genetics:  
A;Gene: apoA-I  
A;Introns: 15/1; 66/2  
C;Superfamily: apolipoprotein A-I  
C;Keywords: cholesterol metabolism; HDL; intestine; lipid binding; lipid metabolism; li  
F;95/Region: ochre stop codon  
Query Match 56.1%; Score 37; DB 2; Length 265;  
Best Local Similarity 61.5%; Pred. No. 46;  
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;  
QY 1 KLVPFATLH 13  
Db 163 KLSPLAEELRDL 175  
RESULT 35  
C96752  
probable DNA-binding protein P28P22.7 [imported] - Arabidopsis thaliana  
C;Species: Arabidopsis thaliana (mouse-ear cress)  
C;Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 09-Jul-2004  
C;Accession: C96752  
R;Theologas, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso  
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.  
ansen, N.F.; Hughes, B.; Huizar, L.

Nature 408, 816-820, 2000  
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.  
C:Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maiti, R.; Marziani,  
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,  
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
A:Title: Sequence and analysis of chromosome 1 of the plant *Arabidopsis*.  
A:Reference number: A86141; MUID:21016719; PMID:11130712  
A:Accession: C96752  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-289 <STO>  
A:Cross-references: UNIPROT:Q9CA18; GB:AE005173; NID:g6648154; PIDN:AAF21154.1; GSPDB:GN  
C:Genetics:  
A:Gene: F28P22.7  
A:Map position: 1

Query Match 56.1%; Score 37; DB 2; Length 289;  
Best Local Similarity 63.6%; Pred. No. 51;  
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 KLVPPFATELHE 11  
||: |||: ||  
Db 266 KLVVIAEMHE 276

RESULT 36  
S35642  
RAD51 protein homolog - chicken  
C:Species: Gallus gallus (chicken)  
C:Date: 10-Dec-1993 #sequence\_revision 10-Nov-1995 #text\_change 09-Jul-2004  
C:Accession: S35642  
R:Bezzubova, O.; Shinohara, A.; Mueller, R.G.; Ogawa, H.; Buerstedde, J.M.  
Nucleic Acids Res. 21, 1577-1580, 1993  
A:Title: A chicken RAD51 homolog is expressed at high levels in lymphoid and reproduc  
A:Reference number: S35642; MUID:93241937; PMID:8479908  
A:Accession: S35642  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-339 <BEZ>  
A:Cross-references: UNIPROT:P37383; GB:S59426; GB:D09655; NID:g299818; PIDN:AAB26354.1;  
C:Superfamily: yeast DNA repair protein RAD51  
C:Keywords: DNA binding

Query Match 56.1%; Score 37; DB 2; Length 339;  
Best Local Similarity 56.2%; Pred. No. 60;  
Matches 9; Conservative 1; Mismatches 2; Indels 4; Gaps 1;

QY 1 KLVPPF----ATELHER 12  
||||| |||: ||  
Db 80 KLVPMGFTTATEFHQR 95

RESULT 37  
I58295  
RAD51 protein homolog - human  
N:Alternate names: recombination protein recA homolog  
C:Species: Homo sapiens (man)  
C:Date: 31-May-1996 #sequence\_revision 31-May-1996 #text\_change 09-Jul-2004  
C:Accession: I58295; S36161; S37673; S35632  
R:Yoshimura, Y.; Morita, T.; Yamamoto, A.; Matsushiro, A.  
Nucleic Acids Res. 21, 1665, 1993  
A:Title: Cloning and sequence of the human RecA-like gene cDNA.  
A:Reference number: I58295; MUID:93241950; PMID:8479919  
A:Accession: I58295  
A:Status: nucleic acid sequence not shown; translated from GB/EMBL/DDBU  
A:Molecule type: mRNA  
A:Residues: 1-339 <RES>  
A:Cross-references: UNIPROT:Q06609; GB:D14134; NID:g285976; PIDN:HAA03189.1; PID:g285977  
R:Shinohara, A.; Ogawa, H.; Matsuda, Y.; Ushio, N.; Ikeo, K.; Ogawa, T.  
Nature Genet. 4, 239-243, 1993  
A:Title: Cloning of human, mouse and fission yeast recombination genes homologous to RAD  
A:Reference number: S36159; MUID:93364417; PMID:8358431

A:Accession: S36161  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-339 <SHI>  
R:Shinohara, A.; Ogawa, H.; Matsuda, Y.; Ushio, N.; Ikeo, K.; Ogawa, T.  
submitted to the EMBL Data Library, December 1992  
A:Reference number: S37672  
A:Accession: S37673  
A:Molecule type: mRNA  
A:Residues: 1-312,'Q' 314-339 <SH2>  
A:Cross-references: GB:D13804; NID:g397826; PIDN:BAA02962.1; PID:g397827  
C:Genetics:  
A:Gene: GDB:RECA  
A:Cross-references: GDB:138758; OMIM:179617  
A:Map position: 15q15.1-15q15.1  
C:Superfamily: yeast DNA repair protein RAD51  
C:Keywords: DNA recombination; DNA repair; nucleotide binding; P-loop  
F;127-134/Region: nucleotide-binding motif A (P-loop)  
F;218-222/Region: nucleotide-binding motif B

Query Match 56.1%; Score 37; DB 2; Length 339;  
Best Local Similarity 56.2%; Pred. No. 60;  
Matches 9; Conservative 1; Mismatches 2; Indels 4; Gaps 1;

QY 1 KLVPPF----ATELHER 12  
||||| |||: ||  
Db 80 KLVPMGFTTATEFHQR 95

RESULT 38  
A48221  
RAD51 protein homolog - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 21-Jan-1994 #sequence\_revision 18-Nov-1994 #text\_change 09-Jul-2004  
C:Accession: A48221; S36160  
R:Morita, T.; Yoshimura, Y.; Yamamoto, A.; Murata, K.; Mori, M.; Yamamoto, H.; Matsushiro,  
Proc. Natl. Acad. Sci. U.S.A. 90, 6577-6580, 1993  
A:Title: A mouse homolog of the *Escherichia coli* recA and *Saccharomyces cerevisiae* RAD5  
A:Reference number: A48221; MUID:93342035; PMID:8341671  
A:Accession: A48221  
A:Status: preliminary  
A:Molecule type: nucleic acid  
A:Residues: 1-339 <MOR>  
A:Cross-references: UNIPROT:Q08297; GB:D13473; NID:g407348; PIDN:BAA02718.1; PID:g40734  
A:Experimental source: 129/Sv, testes  
A:Note: sequence extracted from NCBI backbone (NCBIN:135814, NCBIP:135815)  
R:Shinohara, A.; Ogawa, H.; Matsuda, Y.; Ushio, N.; Ikeo, K.; Ogawa, T.  
Nature Genet. 4, 239-243, 1993  
A:Title: Cloning of human, mouse and fission yeast recombination genes homologous to RA  
A:Reference number: S36159; MUID:93364417; PMID:8358431  
A:Accession: S36160  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-339 <SHI>  
A:Cross-references: EMBL:D13803; NID:g303702; PIDN:BAA02961.1; PID:g397831  
C:Superfamily: yeast DNA repair protein RAD51  
C:Keywords: DNA binding

Query Match 56.1%; Score 37; DB 2; Length 339;  
Best Local Similarity 56.2%; Pred. No. 60;  
Matches 9; Conservative 1; Mismatches 2; Indels 4; Gaps 1;

QY 1 KLVPPF----ATELHER 12  
||||| |||: ||  
Db 80 KLVPMGFTTATEFHQR 95

RESULT 39  
S45881  
probable purine nucleotide-binding protein YBR025c - yeast (*Saccharomyces cerevisiae*)  
N:Alternate names: protein YBR0309  
C:Species: *Saccharomyces cerevisiae*  
C:Date: 26-Aug-1994 #sequence\_revision 09-Sep-1994 #text\_change 09-Jul-2004



C;Accession: S45881; S46556  
 R;Grivell, L.A.; de Haan, M.; Maat, M.J.; Smits, P.H.M.  
 submitted to the Protein Sequence Database, August 1994  
 A;Reference number: S45875  
 A;Accession: S45881  
 A;Molecule type: DNA  
 A;Residues: 1-394 <GRI>  
 A;Cross-references: UNIPROT:P38219; EMBL:Z35894; NID:G536233; PIDN:CAA84967.1; PID:G536233  
 A;Experimental source: strain S288C  
 R;Smits, P.H.M.; de Haan, M.; Maat, C.; Grivell, L.A.  
 Yeast 10(Suppl.A), S75-S80, 1994  
 A;Title: The complete sequence of a 33 kb fragment on the right arm of chromosome II from *Schizosaccharomyces pombe* and a homologue of the SCO1 gene.  
 A;Identified genes and a homologue of the SCO1 gene.  
 A;Reference number: S46551; MUID:94378725; PMID:8091864  
 A;Accession: S46556  
 A;Status: nucleic acid sequence not shown; translation not shown  
 A;Molecule type: DNA  
 A;Residues: 1-394 <SMI>  
 A;Cross-references: EMBL:X76078; NID:G498748; PIDN:CAA53682.1; PID:G498754  
 A;Experimental source: strain S288C  
 A;Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1993  
 C;Genetics:  
 A;Cross-references: SGD:S0000229  
 A;Map position: 2R  
 C;Superfamily: yeast probable purine nucleotide-binding protein YBR025c  
 C;Keywords: ATP; P-loop; purine nucleotide binding  
 F;27-34/Region: nucleotide-binding motif A (P-loop)  
 F;33/Binding site: ATP/GTP (Lys) #status predicted

Query Match 56.1%; Score 37; DB 2; Length 394;  
 Best Local Similarity 50.0%; Pred. No. 71;  
 Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 2 LVPFATLHERL 13  
 :||: ||||  
 Db 262 IIPFVSLEERL 273

## RESULT 40

C83347  
 probable dipeptidase precursor PA2393 [imported] - Pseudomonas aeruginosa (strain PA01)  
 C;Species: Pseudomonas aeruginosa  
 C;Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 09-Jul-2004  
 C;Accession: C83347  
 R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Br  
 adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,  
 ; Lory, S.; Olson, M.V.  
 Nature 406, 959-964, 2000  
 A;Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho  
 A;Reference number: A82950; MUID:20437337; PMID:10984043  
 A;Accession: C83347  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-448 <STO>  
 A;Cross-references: UNIPROT:Q91187; GB:AE004665; GB:AE004091; NID:G9948426; PIDN:AAG0578  
 A;Experimental source: strain PA01  
 C;Genetics:  
 C;Superfamily: membrane dipeptidase

Query Match 56.1%; Score 37; DB 2; Length 448;  
 Best Local Similarity 53.8%; Pred. No. 81;  
 Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 1 KLVFPFATLHERL 13  
 :||: |||||  
 Db 41 KVMKHAELHERI 53

## RESULT 41

T39621  
 peptidyl prolyl cis trans isomerase - fission yeast (*Schizosaccharomyces pombe*)  
 C;Species: Schizosaccharomyces pombe

C;Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 09-Jul-2004  
 C;Accession: T39621  
 R;Wood, V.; Rajadream, M.A.; Barrrell, B.G.; Skelton, J.; Churcher, C.M.  
 submitted to the EMBL Data Library, March 1998  
 A;Reference number: Z21843  
 A;Accession: T39621  
 A;Status: preliminary; translated from GB/EMBL/DDBU  
 A;Molecule type: DNA  
 A;Residues: 1-463 <WOO>  
 A;Cross-references: UNIPROT:O42941; EMBL:AL022104; PIDN:CAA17903.1; GSPDB:GN00067; SPDB  
 A;Experimental source: strain 972h-; cosmid c16H5  
 C;Genetics:  
 A;Gene: SPDB:SPBC16H5.05c  
 A;Map position: 2  
 A;Introns: 2/3; 12/2; 47/2

Query Match 56.1%; Score 37; DB 2; Length 463;  
 Best Local Similarity 70.0%; Pred. No. 84;  
 Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 4 PFATLHERL 13  
 |||||  
 Db 82 PFAVETHPL 91

## RESULT 42

S37664  
 peplomeric polyprotein precursor - avian infectious bronchitis virus (strain D1466) (fr  
 N;Contains: E2 glycoprotein subunit S2  
 C;Species: avian infectious bronchitis virus, IBV  
 A;Variety: strain D1466  
 C;Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 09-Jul-2004  
 C;Accession: S37664  
 R;Kusters, J.G.; Jager, E.J.; Niesters, H.G.M.; van der Zeijst, B.A.M.  
 Vaccine 8, 605-608, 1990  
 A;Title: Sequence evidence for RNA recombination in field isolates of avian coronavirus  
 A;Reference number: S37663; MUID:91205880; PMID:1708184  
 A;Accession: S37664  
 A;Molecule type: genomic RNA  
 A;Residues: 1-621 <KUS>  
 A;Cross-references: UNIPROT:Q66196; EMBL:X58001; NID:G58986; PIDN:CAA41065.1; PID:G58986  
 C;Superfamily: coronavirus E2 glycoprotein  
 C;Keywords: glycoprotein; peplomer protein; spike protein  
 F;1-5/Product: E2 glycoprotein subunit S1 (fragment) #status predicted <GSI>  
 F;6-621/Product: E2 glycoprotein subunit S2 #status predicted <GS2>  
 F;10,47,59,137,144,415,447,482,506,519,542/Binding site: carbohydrate (Asn) (covalent)

Query Match 56.1%; Score 37; DB 2; Length 621;  
 Best Local Similarity 54.5%; Pred. No. 1.2e+02;  
 Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 3 VPFATLHERL 13  
 :|||: ||  
 Db 237 IIPFATLQARI 247

## RESULT 43

JC4510  
 pullulanase (EC 3.2.1.41) precursor - yeast (*Lipomyces kononenkoae*)  
 N;Alternate names: LKA1 protein; raw starch-degrading amylase  
 N;Contains: alpha-dextrin endo-1,6-alpha-glucosidase (EC 3.2.1.41)  
 C;Species: *Lipomyces kononenkoae*  
 C;Date: 07-Feb-1996 #sequence\_revision 23-Aug-1996 #text\_change 09-Jul-2004  
 C;Accession: JC4510; PC4116  
 R;Steyn, A.J.C.; Marmur, J.; Pretorius, I.S.  
 Gene 166, 65-71, 1995  
 A;Title: Cloning, sequence analysis and expression in yeasts of a cDNA containing a Lip  
 A;Reference number: JC4510; MUID:96105202; PMID:8529895  
 A;Accession: JC4510  
 A;Molecule type: mRNA  
 A;Residues: 1-624 <STE>  
 A;Cross-references: UNIPROT:Q01117; GB:U30376; NID:g1173536; PIDN:AAC49622.1; PID:g1173  
 A;Experimental source: strain IGC4052B

A;Accession: PC4116  
A;Molecule type: protein  
A;Residues: 29-44 <ST2>  
A;Experimental source: IGC4052B  
C;Genetics:  
A;Gene: LKA1  
C;Function:  
A;Description: hydrolyzes raw starch, degrading both alpha-1,4 and alpha-1,6 linkages  
A;Pathway: glycogen/starch degradation  
C;Superfamily: Lipomycetes alpha-amylase; alpha-amylase core homology; glucoamylase starch  
C;Keywords: extracellular protein; glycoprotein; glycosidase; hydrolase; polysaccharide  
F;1-28/Domain: signal sequence #status predicted <SIG>  
F;29-624/Product: alpha-amylase #status predicted <MAT>  
F;48-141/Domain: glucoamylase starch-binding domain homology <SBD>  
F;320-447/Domain: alpha-amylase core homology <AMY>  
F;177-185,297-311,387-430,587-622/Disulfide bonds: #status predicted  
F;304,344/Binding site: carbohydrate (Asn) (covalent) #status predicted  
F;357,377,444/Active site: His, Glu, Asp #status predicted

Query Match 56.1%; Score 37; DB 1; Length 624;  
Best Local Similarity 63.6%; Pred. No. 1.2e+02; Mismatches 3; Indels 0; Gaps 0;  
Matches 7; Conservative 1;

QY 2 LVFPATELHER 12  
: : ||||| :  
Db 247. LIALATELHNR 257

RESULT 44  
S37663  
peplomeric polyprotein precursor - avian infectious bronchitis virus (strain D207) (frag  
N;Contains: E2 glycoprotein subunit S2  
C;Species: avian infectious bronchitis virus, IBV  
A;Variety: strain D207  
C;Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 09-Jul-2004  
C;Accession: S37663  
R;Kusters, J.G.; Jager, E.J.; Niesters, H.G.M.; van der Zeijst, B.A.M.  
Vaccine 8, 605-608, 1990  
A;Title: Sequence evidence for RNA recombination in field isolates of avian coronavirus  
A;Reference number: S37663; MUID:91205880; PMID:1708184  
A;Accession: S37663  
A;Molecule type: genomic RNA  
A;Residues: 1-630 <KUS>  
A;Cross-references: UNIPROT:Q661197; EMBL:X58003; NID:958988; PIDN:CAA41067.1; PID:958989  
C;Superfamily: coronavirus E2 glycoprotein  
C;Keywords: glycoprotein; peplomer protein; spike protein  
F;1-5/Product: E2 glycoprotein subunit S1 (fragment) #status predicted <GS1>  
F;6-621/Product: E2 glycoprotein subunit S2 #status predicted <GS2>  
F;10,47,59,137,144,415,447,482,506,519,542/Binding site: carbohydrate (Asn) (covalent) #

Query Match 56.1%; Score 37; DB 2; Length 630;  
Best Local Similarity 54.5%; Pred. No. 1.2e+02; Mismatches 3; Indels 0; Gaps 0;  
Matches 6; Conservative 3;

QY 3 VPATELHERL 13  
: ||||| :  
Db 237 IPFATQIARI 247

RESULT 45  
S39630  
multidrug-efflux transport protein B - Pseudomonas aeruginosa  
N;Alternate names: multidrug resistance protein B  
C;Species: Pseudomonas aeruginosa  
C;Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 09-Jul-2004  
C;Accession: S39630  
R;Poole, K.; Heinrichs, D.E.; Neshat, S.  
Mol. Microbiol. 10, 529-544, 1993  
A;Title: Cloning and sequence analysis of an EnvCD homologue in Pseudomonas aeruginosa:  
A;Reference number: S39629; MUID:95058196; PMID:7968531  
A;Accession: S39630  
A;Status: preliminary  
A;Molecule type: DNA

A;Residues: 1-1046 <POO>  
A;Cross-references: UNIPROT:P52002; GB:L11616; NID:9438852; PIDN:AAA74437.1; PID:943885  
A;Experimental source: strain CD10  
C;Genetics:  
A;Gene: mexB  
C;Function:  
A;Description: probably involved in secretion of the siderophore pyoverdine  
A;Note: inducible under conditions of iron limitation  
C;Superfamily: acriflavin resistance protein  
C;Keywords: transmembrane protein; transport protein  
F;14-30/Domain: transmembrane #status predicted <TM1>  
F;343-359/Domain: transmembrane #status predicted <TM2>  
F;370-386/Domain: transmembrane #status predicted <TM3>  
F;397-413/Domain: transmembrane #status predicted <TM4>  
F;442-458/Domain: transmembrane #status predicted <TM5>  
F;472-488/Domain: transmembrane #status predicted <TM6>  
F;541-557/Domain: transmembrane #status predicted <TM7>  
F;874-890/Domain: transmembrane #status predicted <TM8>  
F;898-914/Domain: transmembrane #status predicted <TM9>  
F;973-989/Domain: transmembrane #status predicted <TM10>  
F;1014-1030/Domain: transmembrane #status predicted <TM11>

Query Match 56.1%; Score 37; DB 2; Length 1046;  
Best Local Similarity 63.6%; Pred. No. 2e+02; Mismatches 2; Indels 0; Gaps 0;  
Matches 7; Conservative 2;

QY 2 LVFPATELHER 12  
: : ||||| :  
Db 944 IVEFAKELHEQ 954

Search completed: September 1, 2005, 16:41:28  
Job time : 42 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

# OM protein - protein search, using sw model

Run on: September 1, 2005, 16:35:05 ; Search time 162 Seconds  
(without alignments)  
31.603 Million call updates/sec

Title: US-09-993-366-1

Perfect score: 66

Sequence: 1 KLVPPATELHERL 13

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1774312 seqs, 393823214 residues

Total number of hits satisfying chosen parameters: 1774312

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

## Database : Published Applications AA:\*

```
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US05_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US10E_PUBCOMB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
19: /cgn2_6/ptodata/1/pubpaa/US11A_PUBCOMB.pep.*
20: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
21: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
22: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	66	100.0	13	10	US-09-991-809-2
2	66	100.0	85	18	US-10-731-875A-46
3	66	100.0	85	18	US-10-731-875A-47
4	66	100.0	85	18	US-10-731-875A-48
5	66	100.0	85	18	US-10-731-875A-49
6	66	100.0	382	9	US-09-800-729-206
7	66	100.0	382	9	US-09-987-107-37
8	66	100.0	382	18	US-10-987-454-28
9	66	100.0	382	19	US-11-017-037-37
10	66	100.0	382	20	US-11-017-059-37
11	66	100.0	396	9	US-09-800-729-207

66	100.0	396	9	US-09-987-107-33	Sequence 33, Appl
66	100.0	396	10	US-09-802-640-16	Sequence 16, Appl
66	100.0	396	15	US-10-403-902A-16	Sequence 16, Appl
66	100.0	396	17	US-10-741-600-1631	Sequence 1631, Ap
66	100.0	396	17	US-10-741-600-1632	Sequence 1632, Ap
66	100.0	396	18	US-10-987-454-7	Sequence 7, Appl
66	100.0	396	19	US-11-017-037-33	Sequence 33, Appl
66	100.0	396	20	US-11-017-059-33	Sequence 36, Appl
66	100.0	401	9	US-09-987-107-36	Sequence 36, Appl
66	100.0	401	18	US-10-987-454-29	Sequence 29, Appl
66	100.0	401	18	US-10-987-454-32	Sequence 32, Appl
66	100.0	401	19	US-11-017-037-35	Sequence 36, Appl
66	100.0	401	20	US-11-017-059-36	Sequence 36, Appl
66	100.0	429	9	US-09-987-107-34	Sequence 34, Appl
66	100.0	429	18	US-10-987-454-30	Sequence 30, Appl
66	100.0	429	19	US-11-017-037-34	Sequence 34, Appl
66	100.0	429	20	US-11-017-059-34	Sequence 34, Appl
45	68.2	85	18	US-10-731-875A-45	Sequence 45, Appl
41	62.1	258	14	US-10-316-253-279	Sequence 279, App
41	62.1	258	14	US-10-316-253-281	Sequence 281, App
41	62.1	258	14	US-10-316-253-283	Sequence 283, App
41	62.1	472	15	US-10-417-700A-79	Sequence 79, Appl
40	60.6	166	16	US-10-767-701-61044	Sequence 61044, A
35	60.6	237	16	US-10-425-115-242838	Sequence 242838, A
40	60.6	239	16	US-10-437-963-120340	Sequence 120340, A
37	60.6	266	9	US-09-987-107-20	Sequence 20, Appl
38	60.6	266	18	US-10-987-454-17	Sequence 17, Appl
39	60.6	266	19	US-11-017-037-20	Sequence 20, Appl
40	60.6	266	20	US-11-017-059-20	Sequence 20, Appl
41	60.6	1809	16	US-10-437-963-195972	Sequence 195972, A
42	59.1	86	16	US-10-437-963-177544	Sequence 177544, A
43	59.1	117	16	US-10-437-963-152800	Sequence 152800, A
44	59.1	143	16	US-10-425-115-336871	Sequence 336871, A
45	59.1	220	9	US-09-815-242-10585	Sequence 10585, A
46	59.1	220	15	US-10-282-122A-57019	Sequence 57019, A
47	59.1	230	15	US-10-425-114-3637	Sequence 3637, A
48	59.1	397	15	US-10-369-493-3720	Sequence 3720, Ap
49	59.1	397	15	US-10-369-493-3721	Sequence 3721, Ap
50	59.1	553	16	US-10-437-963-114148	Sequence 114148, A
51	59.1	2796	9	US-09-870-759-114	Sequence 114, App
52	59.1	2796	10	US-09-751-708A-114	Sequence 114, App
53	59.1	2796	16	US-10-741-191-31	Sequence 31, Appl
54	59.1	2796	16	US-10-742-350-31	Sequence 31, Appl
55	59.1	2796	16	US-10-428-817A-110	Sequence 110, App
56	59.1	2796	17	US-10-937-758A-91	Sequence 91, Appl
57	59.1	3069	9	US-09-712-363-246	Sequence 246, App
58	59.1	3069	16	US-10-741-191-23	Sequence 23, Appl
59	59.1	3069	16	US-10-742-350-23	Sequence 23, Appl
60	59.1	3076	16	US-10-741-191-24	Sequence 24, Appl
61	59.1	3076	16	US-10-742-350-24	Sequence 24, Appl
62	57.6	80	15	US-10-424-599-155072	Sequence 155072, A
63	57.6	261	16	US-10-425-115-348464	Sequence 348464, A
64	57.6	262	14	US-10-032-189-79	Sequence 79, Appl
65	57.6	277	15	US-10-038-854-235	Sequence 225, App
66	57.6	277	15	US-10-038-854-232	Sequence 232, App
67	57.6	312	16	US-10-425-115-348462	Sequence 348462, A
68	57.6	314	14	US-10-146-772-10	Sequence 10, Appl
69	57.6	314	15	US-10-241-742-10	Sequence 10, Appl
70	57.6	314	15	US-10-440-523-10	Sequence 10, Appl
71	57.6	314	15	US-10-440-503-10	Sequence 10, Appl
72	57.6	314	15	US-10-461-925-10	Sequence 10, Appl
73	57.6	391	9	US-09-987-107-38	Sequence 38, Appl
74	57.6	391	14	US-10-316-253-238	Sequence 238, App
75	57.6	391	14	US-10-316-253-240	Sequence 240, App
76	57.6	391	14	US-10-316-253-242	Sequence 242, App
77	57.6	391	18	US-10-987-454-25	Sequence 25, Appl
78	57.6	391	19	US-11-017-037-38	Sequence 38, Appl
79	57.6	391	20	US-11-017-059-38	Sequence 38, Appl
80	57.6	405	15	US-10-369-493-11584	Sequence 11584, A
81	57.6	405	15	US-10-369-493-14239	Sequence 14239, A
82	57.6	405	15	US-10-369-493-14854	Sequence 14854, A
83	57.6	405	15	US-10-369-493-15043	Sequence 15043, A
84	57.6	405	15	US-10-369-493-23391	Sequence 23391, A

85 38 57.6 428 15 US-10-369-493-17483 Sequence 17483, A  
86 38 57.6 428 15 US-10-282-122A-45770 Sequence 45770, A  
87 38 57.6 429 15 US-10-282-122A-45455 Sequence 45455, A  
88 38 57.6 473 15 US-10-389-566-1358 Sequence 1358, Ap  
89 38 57.6 516 15 US-10-425-114-63544 Sequence 63544, A  
90 38 57.6 542 16 US-10-425-115-331388 Sequence 331388,  
91 38 57.6 652 14 US-10-156-761-12996 Sequence 12996, A  
92 38 57.6 669 17 US-10-892-544-2 Sequence 2, Appli  
93 38 57.6 107 18 US-10-731-875A-38 Sequence 38, Appl  
94 37 56.1 107 18 US-10-731-875A-39 Sequence 39, Appl  
95 37 56.1 124 16 US-10-425-115-244210 Sequence 244210,  
96 37 56.1 166 15 US-10-389-674-44 Sequence 44, Appl  
97 37 56.1 166 15 US-10-389-674-56 Sequence 56, Appl  
98 37 56.1 166 15 US-10-389-674-65 Sequence 65, Appl  
99 37 56.1 170 16 US-10-437-963-155938 Sequence 155938,  
100 37 56.1 239 18 US-10-987-454-14 Sequence 14, Appl

## ALIGNMENTS

RESULT 1  
US-09-991-809-2  
; Sequence 2, Application US/09991809  
; Publication No. US20030100014A1  
; GENERAL INFORMATION:  
; APPLICANT: Jackowski, George  
; TITLE OF INVENTION: Apolipoprotein Biopolymer Markers Predictive of Type II Diabetes  
; FILE REFERENCE: 2132.111  
; CURRENT APPLICATION NUMBER: US/09/991,809  
; CURRENT FILING DATE: 2001-11-23  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 2  
; LENGTH: 13  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-991-809-2

Query Match 100.0%; Score 66; DB 10; Length 13;  
Best Local Similarity 100.0%; Pred. No. 3.2e-05;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KLVPPFATELHERL 13  
Db 1 KLVPPFATELHERL 13

RESULT 2  
US-10-731-875A-46  
; Sequence 46, Application US/10731875A  
; Publication No. US20050149269A1  
; GENERAL INFORMATION:  
; APPLICANT: Applera Corporation  
; APPLICANT: Thomas, Paul  
; APPLICANT: Kejariwal, Anish  
; APPLICANT: Campbell, Michael  
; APPLICANT: Mi, Huaiyu  
; APPLICANT: Diemer, Karen  
; APPLICANT: Guo, Nan  
; APPLICANT: Ladunga, Istvan  
; APPLICANT: Lazareva, Betty  
; APPLICANT: Muruganujan, Anushya  
; APPLICANT: Rabkin, Steven  
; APPLICANT: Vandergriff, Jody  
; TITLE OF INVENTION: Browseable Database For Biological Use  
; FILE REFERENCE: 9692-000029  
; CURRENT APPLICATION NUMBER: US/10/731,875A  
; CURRENT FILING DATE: 2003-12-09  
; NUMBER OF SEQ ID NOS: 63  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 46

; LENGTH: 85  
; TYPE: PRT  
; ORGANISM: Unknown  
; FEATURE:  
; OTHER INFORMATION: Exemplary Protein Sequence.  
US-10-731-875A-46

Query Match 100.0%; Score 66; DB 18; Length 85;  
Best Local Similarity 100.0%; Pred. No. 0.00026;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KLVPPFATELHERL 13  
Db 28 KLVPPFATELHERL 40

RESULT 3  
US-10-731-875A-47  
; Sequence 47, Application US/10731875A  
; Publication No. US20050149269A1  
; GENERAL INFORMATION:  
; APPLICANT: Applera Corporation  
; APPLICANT: Thomas, Paul  
; APPLICANT: Kejariwal, Anish  
; APPLICANT: Campbell, Michael  
; APPLICANT: Mi, Huaiyu  
; APPLICANT: Diemer, Karen  
; APPLICANT: Guo, Nan  
; APPLICANT: Ladunga, Istvan  
; APPLICANT: Lazareva, Betty  
; APPLICANT: Muruganujan, Anushya  
; APPLICANT: Rabkin, Steven  
; APPLICANT: Vandergriff, Jody  
; APPLICANT: Doremieux, Olivier  
; TITLE OF INVENTION: Browseable Database For Biological Use  
; FILE REFERENCE: 9692-000029  
; CURRENT APPLICATION NUMBER: US/10/731,875A  
; CURRENT FILING DATE: 2003-12-09  
; NUMBER OF SEQ ID NOS: 63  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 47  
; LENGTH: 85  
; TYPE: PRT  
; ORGANISM: Unknown  
; FEATURE:  
; OTHER INFORMATION: Exemplary Protein Sequence.  
US-10-731-875A-47

Query Match 100.0%; Score 66; DB 18; Length 85;  
Best Local Similarity 100.0%; Pred. No. 0.00026;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KLVPPFATELHERL 13  
Db 28 KLVPPFATELHERL 40

RESULT 4  
US-10-731-875A-48  
; Sequence 48, Application US/10731875A  
; Publication No. US20050149269A1  
; GENERAL INFORMATION:  
; APPLICANT: Applera Corporation  
; APPLICANT: Thomas, Paul  
; APPLICANT: Kejariwal, Anish  
; APPLICANT: Campbell, Michael  
; APPLICANT: Mi, Huaiyu  
; APPLICANT: Diemer, Karen  
; APPLICANT: Guo, Nan  
; APPLICANT: Ladunga, Istvan  
; APPLICANT: Lazareva, Betty  
; APPLICANT: Muruganujan, Anushya  
; APPLICANT: Rabkin, Steven

; APPLICANT: Vandergriff, Jody  
; APPLICANT: Doremieux, Olivier  
; TITLE OF INVENTION: Browseable Database For Biological Use  
; FILE REFERENCE: 9692-000029  
; CURRENT APPLICATION NUMBER: US/10/731,875A  
; CURRENT FILING DATE: 2003-12-09  
; NUMBER OF SEQ ID NOS: 63  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 48  
; LENGTH: 85  
; TYPE: PRT  
; ORGANISM: Unknown  
; FEATURE:  
; OTHER INFORMATION: Exemplary Protein Sequence.  
US-10-731-875A-48

Query Match 100.0%; Score 66; DB 18; Length 85;  
Best Local Similarity 100.0%; Pred. No. 0.00026;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLVPPFATLHERL 13  
| | | | | | | | | | | | | | |  
Db 28 KLVPPFATLHERL 40

RESULT 5  
US-10-731-875A-49  
; Sequence 49, Application US/10731875A  
; Publication No. US20050149269A1  
; GENERAL INFORMATION:  
; APPLICANT: Applera Corporation  
; APPLICANT: Thomas, Paul  
; APPLICANT: Kejariwal, Anish  
; APPLICANT: Campbell, Michael  
; APPLICANT: Mi, Huaiyu  
; APPLICANT: Diemer, Karen  
; APPLICANT: Guo, Nan  
; APPLICANT: Ladunga, Istvan  
; APPLICANT: Lazareva, Betty  
; APPLICANT: Muruganujan, Anushya  
; APPLICANT: Rabkin, Steven  
; APPLICANT: Vandergriff, Jody  
; APPLICANT: Doremieux, Olivier  
; TITLE OF INVENTION: Browseable Database For Biological Use  
; FILE REFERENCE: 9692-000029  
; CURRENT APPLICATION NUMBER: US/10/731,875A  
; CURRENT FILING DATE: 2003-12-09  
; NUMBER OF SEQ ID NOS: 63  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 49  
; LENGTH: 85  
; TYPE: PRT  
; ORGANISM: Unknown  
; FEATURE:  
; OTHER INFORMATION: Exemplary Protein Sequence.  
US-10-731-875A-49

Query Match 100.0%; Score 66; DB 18; Length 85;  
Best Local Similarity 100.0%; Pred. No. 0.00026;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLVPPFATLHERL 13  
| | | | | | | | | | | | | | |  
Db 28 KLVPPFATLHERL 40

RESULT 6  
US-09-800-729-206  
; Sequence 206, Application US/09800729  
; Patent No. US2002006819A1  
; GENERAL INFORMATION:  
; APPLICANT: Ni et al.  
; TITLE OF INVENTION: 32 Human secreted proteins

; FILE REFERENCE: P2044PI  
; CURRENT APPLICATION NUMBER: US/09/800,729  
; CURRENT FILING DATE: 2001-03-08  
; PRIOR APPLICATION NUMBER: PCT/US00/26013  
; PRIOR FILING DATE: 2000-09-22  
; PRIOR APPLICATION NUMBER: 60/155,709  
; PRIOR FILING DATE: 1999-09-24  
; NUMBER OF SEQ ID NOS: 217  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 206  
; LENGTH: 382  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-800-729-206

Query Match 100.0%; Score 66; DB 9; Length 382;  
Best Local Similarity 100.0%; Pred. No. 0.0014;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLVPPFATLHERL 13  
| | | | | | | | | | | | | | |  
Db 79 KLVPPFATLHERL 91

RESULT 7  
US-09-987-107-37  
; Sequence 37, Application US/09987107  
; Patent No. US20020156007A1  
; GENERAL INFORMATION:  
; APPLICANT: GRAVERSEN, Jonas  
; APPLICANT: MOESTRUP, Soren  
; TITLE OF INVENTION: APOLIPOPROTEINS ANALOGUES  
; FILE REFERENCE: GRAVERSEN1A  
; CURRENT APPLICATION NUMBER: US/09/987,107  
; CURRENT FILING DATE: 2001-11-13  
; PRIOR APPLICATION NUMBER: US 60/264,022  
; PRIOR FILING DATE: 2001-01-26  
; PRIOR APPLICATION NUMBER: DK PA2001 00057  
; PRIOR FILING DATE: 2001-01-15  
; PRIOR APPLICATION NUMBER: DK PA2000 01682  
; PRIOR FILING DATE: 2000-11-10  
; NUMBER OF SEQ ID NOS: 91  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 37  
; LENGTH: 382  
; TYPE: PRT  
; ORGANISM: Sus scrofa  
US-09-987-107-37

Query Match 100.0%; Score 66; DB 9; Length 382;  
Best Local Similarity 100.0%; Pred. No. 0.0014;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLVPPFATLHERL 13  
| | | | | | | | | | | | | | |  
Db 79 KLVPPFATLHERL 91

RESULT 8  
US-10-987-454-28  
; Sequence 28, Application US/10987454  
; Publication No. US20050172359A1  
; GENERAL INFORMATION:  
; APPLICANT: Reid, Alexandra  
; APPLICANT: Moloney, Maurice  
; TITLE OF INVENTION: METHODS FOR THE PRODUCTION OF APOLIPOPROTEINS IN TRANSGENIC  
; TITLE OF INVENTION: PLANTS  
; FILE REFERENCE: 9369-311  
; CURRENT APPLICATION NUMBER: US/10/987,454  
; CURRENT FILING DATE: 2004-11-15  
; PRIOR APPLICATION NUMBER: US 60/519,606  
; PRIOR FILING DATE: 2003-11-14  
; PRIOR APPLICATION NUMBER: US 60/579,733



Query Match 100.0%; Score 66; DB 9; Length 396;  
Best Local Similarity 100.0%; Pred. No. 0.0014;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLVPFATELHERL 13  
Db 79 KLVPFATELHERL 91

## RESULT 13

US-09-802-640-16  
Sequence 16, Application US/09802640  
Publication No. US20030036057A1

GENERAL INFORMATION:  
APPLICANT: Braun, Andreas  
APPLICANT: Bansal Aruna  
APPLICANT: Kleyn Patrick  
TITLE OF INVENTION: GENES AND POLYMORPHISMS ASSOCIATED WITH  
TITLE OF INVENTION: CARDIOVASCULAR DISEASE AND THEIR USE  
FILE REFERENCE: 24736-2048  
CURRENT APPLICATION NUMBER: US/09/802,640  
CURRENT FILING DATE: 2001-03-09  
NUMBER OF SEQ ID NOS: 122  
SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 16  
LENGTH: 396  
TYPE: PRT  
ORGANISM: Homo sapien

US-09-802-640-16

Query Match 100.0%; Score 66; DB 10; Length 396;  
Best Local Similarity 100.0%; Pred. No. 0.0014;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLVPFATELHERL 13  
Db 79 KLVPFATELHERL 91

## RESULT 14

US-10-403-902A-16  
Sequence 16, Application US/10403902A  
Publication No. US20030224418A1

GENERAL INFORMATION:  
APPLICANT: Braun, Andreas  
APPLICANT: Bansal, Aruna  
APPLICANT: Kleyn, Patrick  
TITLE OF INVENTION: GENES AND POLYMORPHISMS ASSOCIATED WITH  
TITLE OF INVENTION: CARDIOVASCULAR DISEASE AND THEIR USE  
FILE REFERENCE: 24736-2048B  
CURRENT APPLICATION NUMBER: US/10/403,902A  
CURRENT FILING DATE: 2003-07-21  
PRIOR APPLICATION NUMBER: 09/802,640  
PRIOR FILING DATE: 2001-03-09  
NUMBER OF SEQ ID NOS: 122  
SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 16  
LENGTH: 396  
TYPE: PRT  
ORGANISM: Homo sapien

US-10-403-902A-16

Query Match 100.0%; Score 66; DB 15; Length 396;  
Best Local Similarity 100.0%; Pred. No. 0.0014;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLVPFATELHERL 13  
Db 79 KLVPFATELHERL 91

## RESULT 15

US-10-741-600-1631

Sequence 1631, Application US/10741600  
Publication No. US20050026169A1  
GENERAL INFORMATION:  
APPLICANT: CARGILL, Michele et al.  
TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH  
FILE REFERENCE: CL001499  
CURRENT APPLICATION NUMBER: US/10/741,600  
CURRENT FILING DATE: 2003-12-22  
NUMBER OF SEQ ID NOS: 73997  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 1631  
LENGTH: 396  
TYPE: PRT  
ORGANISM: Homo sapiens

US-10-741-600-1631

Query Match 100.0%; Score 66; DB 17; Length 396;  
Best Local Similarity 100.0%; Pred. No. 0.0014;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLVPFATELHERL 13  
Db 79 KLVPFATELHERL 91

## RESULT 16

US-10-741-600-1632  
Sequence 1632, Application US/10741600  
Publication No. US20050026169A1

GENERAL INFORMATION:  
APPLICANT: CARGILL, Michele et al.  
TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH  
FILE REFERENCE: CL001499  
CURRENT APPLICATION NUMBER: US/10/741,600  
CURRENT FILING DATE: 2003-12-22  
NUMBER OF SEQ ID NOS: 73997  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 1632  
LENGTH: 396  
TYPE: PRT  
ORGANISM: Homo sapiens

US-10-741-600-1632

Query Match 100.0%; Score 66; DB 17; Length 396;  
Best Local Similarity 100.0%; Pred. No. 0.0014;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLVPFATELHERL 13  
Db 79 KLVPFATELHERL 91

## RESULT 17

US-10-987-454-7  
Sequence 7, Application US/10987454  
Publication No. US20050172359A1

GENERAL INFORMATION:  
APPLICANT: Reid, Alexandra  
APPLICANT: Moloney, Maurice  
TITLE OF INVENTION: METHODS FOR THE PRODUCTION OF APOLIPOPROTEINS IN TRANSGENIC  
TITLE OF INVENTION: PLANTS  
FILE REFERENCE: 9369-311  
CURRENT APPLICATION NUMBER: US/10/987,454  
CURRENT FILING DATE: 2004-11-15  
PRIOR APPLICATION NUMBER: US 60/519,606  
PRIOR FILING DATE: 2003-11-14  
PRIOR APPLICATION NUMBER: US 60/579,733  
PRIOR FILING DATE: 2004-06-16  
NUMBER OF SEQ ID NOS: 251  
SOFTWARE: PatentIn version 3.3  
SEQ ID NO 7





Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLVFPFATLHERL 13  
Db 63 KLVFPFATLHERL 75

## RESULT 22

US-10-987-454-32  
; Sequence 32, Application US/10987454  
; Publication No. US20050172359A1  
; GENERAL INFORMATION:  
; APPLICANT: Reid, Alexandra  
; APPLICANT: Moloney, Maurice  
; TITLE OF INVENTION: METHODS FOR THE PRODUCTION OF APOLIPOPROTEINS IN TRANSGENIC  
; TITLE OF INVENTION: PLANTS  
; FILE REFERENCE: 9369-311  
; CURRENT APPLICATION NUMBER: US/10/987,454  
; CURRENT FILING DATE: 2004-11-15  
; PRIOR APPLICATION NUMBER: US 60/519,606  
; PRIOR FILING DATE: 2003-11-14  
; PRIOR APPLICATION NUMBER: US 60/579,733  
; PRIOR FILING DATE: 2004-06-16  
; NUMBER OF SEQ ID NOS: 251  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 32  
; LENGTH: 401  
; TYPE: PRT  
; ORGANISM: Papio sp.  
US-10-987-454-32

Query Match 100.0%; Score 66; DB 18; Length 401;  
Best Local Similarity 100.0%; Pred. No. 0.0015;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLVFPFATLHERL 13  
Db 63 KLVFPFATLHERL 75

## RESULT 23

US-11-017-037-36  
; Sequence 36, Application US/11017037  
; Publication No. US20050096277A1  
; GENERAL INFORMATION:  
; APPLICANT: GRAVERSEN, Jonas  
; APPLICANT: MOESTRUP, Soren  
; TITLE OF INVENTION: APOLIPOPROTEINS ANALOGUES  
; FILE REFERENCE: GRAVERSENIA  
; CURRENT APPLICATION NUMBER: US/11/017,037  
; CURRENT FILING DATE: 2004-12-21  
; PRIOR APPLICATION NUMBER: US/09/987,107  
; PRIOR FILING DATE: 2001-11-13  
; PRIOR APPLICATION NUMBER: US 60/264,022  
; PRIOR FILING DATE: 2001-01-26  
; PRIOR APPLICATION NUMBER: DK PA2001 00057  
; PRIOR FILING DATE: 2001-01-15  
; PRIOR APPLICATION NUMBER: DK PA2000 01682  
; PRIOR FILING DATE: 2000-11-10  
; NUMBER OF SEQ ID NOS: 91  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 36  
; LENGTH: 401  
; TYPE: PRT  
; ORGANISM: Papio anubis  
US-11-017-037-36

Query Match 100.0%; Score 66; DB 19; Length 401;  
Best Local Similarity 100.0%; Pred. No. 0.0015;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLVFPFATLHERL 13  
Db 79 KLVFPFATLHERL 91

Db 63 KLVFPFATLHERL 75

## RESULT 24

US-11-017-059-36  
; Sequence 36, Application US/11017059  
; Publication No. US20050142639A1  
; GENERAL INFORMATION:  
; APPLICANT: GRAVERSEN, Jonas  
; APPLICANT: MOESTRUP, Soren  
; TITLE OF INVENTION: APOLIPOPROTEINS ANALOGUES  
; FILE REFERENCE: GRAVERSENIA  
; CURRENT APPLICATION NUMBER: US/11/017,059  
; CURRENT FILING DATE: 2004-12-21  
; PRIOR APPLICATION NUMBER: US/09/987,107  
; PRIOR FILING DATE: 2001-11-13  
; PRIOR APPLICATION NUMBER: US 60/264,022  
; PRIOR FILING DATE: 2001-01-26  
; PRIOR APPLICATION NUMBER: DK PA2001 00057  
; PRIOR FILING DATE: 2001-01-15  
; PRIOR APPLICATION NUMBER: DK PA2000 01682  
; PRIOR FILING DATE: 2000-11-10  
; NUMBER OF SEQ ID NOS: 91  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 36  
; LENGTH: 401  
; TYPE: PRT  
; ORGANISM: Papio anubis  
US-11-017-059-36

Query Match 100.0%; Score 66; DB 20; Length 401;  
Best Local Similarity 100.0%; Pred. No. 0.0015;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLVFPFATLHERL 13  
Db 63 KLVFPFATLHERL 75

## RESULT 25

US-09-987-107-34  
; Sequence 34, Application US/09987107  
; Patent No. US20020156007A1  
; GENERAL INFORMATION:  
; APPLICANT: GRAVERSEN, Jonas  
; APPLICANT: MOESTRUP, Soren  
; TITLE OF INVENTION: APOLIPOPROTEINS ANALOGUES  
; FILE REFERENCE: GRAVERSENIA  
; CURRENT APPLICATION NUMBER: US/09/987,107  
; CURRENT FILING DATE: 2001-11-13  
; PRIOR APPLICATION NUMBER: US 60/264,022  
; PRIOR FILING DATE: 2001-01-26  
; PRIOR APPLICATION NUMBER: DK PA2001 00057  
; PRIOR FILING DATE: 2001-01-15  
; PRIOR APPLICATION NUMBER: DK PA2000 01682  
; PRIOR FILING DATE: 2000-11-10  
; NUMBER OF SEQ ID NOS: 91  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 34  
; LENGTH: 429  
; TYPE: PRT  
; ORGANISM: Macaca fascicularis  
US-09-987-107-34

Query Match 100.0%; Score 66; DB 9; Length 429;  
Best Local Similarity 100.0%; Pred. No. 0.0016;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLVFPFATLHERL 13  
Db 79 KLVFPFATLHERL 91

```

RESULT 26
US-10-987-454-30
; Sequence 30, Application US/10987454
; Publication No. US20050172359A1
; GENERAL INFORMATION:
; APPLICANT: Reid, Alexandra
; APPLICANT: Moloney, Maurice
; TITLE OF INVENTION: METHODS FOR THE PRODUCTION OF APOLIPOPROTEINS IN TRANSGENIC
; TITLE OF INVENTION: PLANTS
; FILE REFERENCE: 9369-311
; CURRENT APPLICATION NUMBER: US/10/987,454
; CURRENT FILING DATE: 2004-11-15
; PRIOR APPLICATION NUMBER: US 60/519,606
; PRIOR FILING DATE: 2003-11-14
; PRIOR APPLICATION NUMBER: US 60/579,733
; PRIOR FILING DATE: 2004-06-16
; NUMBER OF SEQ ID NOS: 251
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 30
; LENGTH: 429
; TYPE: PRT
; ORGANISM: Macaca fascicularis
US-10-987-454-30

```

```

Query Match 100.0%; Score 66; DB 18; Length 429;
Best Local Similarity 100.0%; Pred. No. 0.0016;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 KLVPPFATELHERL 13
Db 79 KLVPPFATELHERL 91

```

```

RESULT 27
US-11-017-037-34
; Sequence 34, Application US/11017037
; Publication No. US20050096277A1
; GENERAL INFORMATION:
; APPLICANT: GRAVERSEN, Jonas
; APPLICANT: MOESTRUP, Soren
; TITLE OF INVENTION: APOLIPOPROTEINS ANALOGUES
; FILE REFERENCE: GRAVERSENIA
; CURRENT APPLICATION NUMBER: US/11/017,037
; CURRENT FILING DATE: 2004-12-21
; PRIOR APPLICATION NUMBER: US/09/987,107
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/264,022
; PRIOR FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: DK PA2001 00057
; PRIOR FILING DATE: 2001-01-15
; PRIOR APPLICATION NUMBER: DK PA2000 01682
; PRIOR FILING DATE: 2000-11-10
; NUMBER OF SEQ ID NOS: 91
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 34
; LENGTH: 429
; TYPE: PRT
; ORGANISM: Macaca fascicularis
US-11-017-037-34

```

```

Query Match 100.0%; Score 66; DB 19; Length 429;
Best Local Similarity 100.0%; Pred. No. 0.0016;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 KLVPPFATELHERL 13
Db 79 KLVPPFATELHERL 91

```

```

RESULT 28
US-11-017-059-34
; Sequence 34, Application US/11017059
; Publication No. US20050142639A1

```

```

; GENERAL INFORMATION:
; APPLICANT: GRAVERSEN, Jonas
; APPLICANT: MOESTRUP, Soren
; TITLE OF INVENTION: APOLIPOPROTEINS ANALOGUES
; FILE REFERENCE: GRAVERSENIA
; CURRENT APPLICATION NUMBER: US/11/017,059
; CURRENT FILING DATE: 2004-12-21
; PRIOR APPLICATION NUMBER: US/09/987,107
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/264,022
; PRIOR FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: DK PA2001 00057
; PRIOR FILING DATE: 2001-01-15
; PRIOR APPLICATION NUMBER: DK PA2000 01682
; PRIOR FILING DATE: 2000-11-10
; NUMBER OF SEQ ID NOS: 91
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 34
; LENGTH: 429
; TYPE: PRT
; ORGANISM: Macaca fascicularis
US-11-017-059-34

```

```

Query Match 100.0%; Score 66; DB 20; Length 429;
Best Local Similarity 100.0%; Pred. No. 0.0016;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 KLVPPFATELHERL 13
Db 79 KLVPPFATELHERL 91

```

```

RESULT 29
US-10-731-875A-45
; Sequence 45, Application US/10731875A
; Publication No. US20050149269A1
; GENERAL INFORMATION:
; APPLICANT: Applera Corporation
; APPLICANT: Thomas, Paul
; APPLICANT: Kejariwal, Anish
; APPLICANT: Campbell, Michael
; APPLICANT: Mi, Huaiyu
; APPLICANT: Diemer, Karen
; APPLICANT: Guo, Nan
; APPLICANT: Ladunga, Istvan
; APPLICANT: Lazareva, Betty
; APPLICANT: Muruganujan, Anushya
; APPLICANT: Rabkin, Steven
; APPLICANT: Vandergrieff, Jody
; APPLICANT: Doremieux, Olivier
; TITLE OF INVENTION: Browsable Database For Biological Use
; FILE REFERENCE: 9692-000029
; CURRENT APPLICATION NUMBER: US/10/731,875A
; CURRENT FILING DATE: 2003-12-09
; NUMBER OF SEQ ID NOS: 63
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 45
; LENGTH: 85
; TYPE: PRT
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: Exemplary Protein Sequence.
US-10-731-875A-45

```

```

Query Match 68.2%; Score 45; DB 18; Length 85;
Best Local Similarity 69.2%; Pred. No. 1.9;
Matches 9; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

```

```

QY 1 KLVPPFATELHERL 13
Db 28 KLVPPFATELQAQL 40

```

```
RESULT 30
US-10-316-253-279
; Sequence 279, Application US/10316253
; Publication No. US20030162706A1
; GENERAL INFORMATION:
; APPLICANT: The Procter & Gamble Company
; APPLICANT: Peters, Kevin
; APPLICANT: Thompson, Larry
; APPLICANT: Wang, Feng
; APPLICANT: Greig, Kenneth
; TITLE OF INVENTION: Angiogenesis Modulating Proteins
; FILE REFERENCE: 8865M
; CURRENT APPLICATION NUMBER: US/10/316,253
; CURRENT FILING DATE: 2002-12-10
; PRIOR APPLICATION NUMBER: US 60/355,295
; PRIOR FILING DATE: 2002-02-08
; NUMBER OF SEQ ID NOS: 308
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 279
; LENGTH: 258
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-10-316-253-279

Query Match      62.1%; Score 41; DB 14; Length 258;
Best Local Similarity 72.7%; Pred. No. 37;
Matches      8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy      1 KLVPPFATLHE 11
Db      140 KLEPLATLHK 150

RESULT 31
US-10-316-253-281
; Sequence 281, Application US/10316253
; Publication No. US20030162706A1
; GENERAL INFORMATION:
; APPLICANT: The Procter & Gamble Company
; APPLICANT: Peters, Kevin
; APPLICANT: Thompson, Larry
; APPLICANT: Wang, Feng
; APPLICANT: Greis, Kenneth
; TITLE OF INVENTION: Angiogenesis Modulating Proteins
; FILE REFERENCE: 8865M
; CURRENT APPLICATION NUMBER: US/10/316,253
; CURRENT FILING DATE: 2002-12-10
; PRIOR APPLICATION NUMBER: US 60/355,295
; PRIOR FILING DATE: 2002-02-08
; NUMBER OF SEQ ID NOS: 308
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 281
; LENGTH: 258
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-10-316-253-281

Query Match      62.1%; Score 41; DB 14; Length 258;
Best Local Similarity 72.7%; Pred. No. 37;
Matches      8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy      1 KLVPPFATLHE 11
Db      140 KLEPLATLHK 150

RESULT 32
US-10-316-253-283
; Sequence 283, Application US/10316253
; Publication No. US20030162706A1
; GENERAL INFORMATION:
; APPLICANT: The Procter & Gamble Company
; APPLICANT: Peters, Kevin
```

```
; APPLICANT: Thompson, Larry
; APPLICANT: Wang, Feng
; APPLICANT: Greis, Kenneth
; TITLE OF INVENTION: Angiogenesis Modulating Proteins
; FILE REFERENCE: 8865M
; CURRENT APPLICATION NUMBER: US/10/316,253
; CURRENT FILING DATE: 2002-12-10
; PRIOR APPLICATION NUMBER: US 60/355,295
; PRIOR FILING DATE: 2002-02-08
; NUMBER OF SEQ ID NOS: 308
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 283
; LENGTH: 258
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-10-316-253-283

Query Match      62.1%; Score 41; DB 14; Length 258;
Best Local Similarity 72.7%; Pred. No. 37;
Matches      8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy      1 KLVPPFATLHE 11
Db      140 KLEPLATLHK 150

RESULT 33
US-10-417-700A-79
; Sequence 79, Application US/10417700A
; Publication No. US20040033581A1
; GENERAL INFORMATION:
; APPLICANT: ECOPIA BIOSCIENCES INC.
; APPLICANT: ZAZOPOULOS, Emmanuel
; APPLICANT: STAFEA, Alfredo
; APPLICANT: FARNET, Chris
; TITLE OF INVENTION: Specialized dual condensation/epimerization domain in non-ribos
; FILE REFERENCE: 3002-14US
; CURRENT APPLICATION NUMBER: US/10/417,700A
; CURRENT FILING DATE: 2003-04-17
; NUMBER OF SEQ ID NOS: 139
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 79
; LENGTH: 472
; TYPE: PRT
; ORGANISM: Streptomyces aizensis strain NRRL B-11277
US-10-417-700A-79

Query Match      62.1%; Score 41; DB 15; Length 472;
Best Local Similarity 72.7%; Pred. No. 72;
Matches      8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy      3 VPPFATLHERL 13
Db      218 VPPFATLHERL 228

RESULT 34
US-10-767-701-61044
; Sequence 61044, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 61044
; LENGTH: 166
```

```
; TYPE: PRT
; ORGANISM: Sorghum bicolor
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(237)
; OTHER INFORMATION: Clone ID: 9300886.pap
US-10-767-701-61044

Query Match      60.6%; Score 40; DB 16; Length 166;
Best Local Similarity 72.7%; Pred. No. 34;
Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 VPFATELHERL 13
   :|||||:|
Db 1 MPFATELTRKL 11

RESULT 35
US-10-425-115-242838
; Sequence 242838, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 242838
; LENGTH: 237
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(237)
; OTHER INFORMATION: unsure at all Xaa locations
; OTHER INFORMATION: Clone ID: MRT4577_15304C.1.pap
US-10-425-115-242838

Query Match      60.6%; Score 40; DB 16; Length 237;
Best Local Similarity 88.9%; Pred. No. 51;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 LVPFATELH 10
   :|||||:|
Db 119 LVPFATEEH 127

RESULT 36
US-10-437-963-120340
; Sequence 120340, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 120340
; LENGTH: 239
; TYPE: PRT
```

```
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(239)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_2346C.1.pap
US-10-437-963-120340

Query Match      60.6%; Score 40; DB 16; Length 239;
Best Local Similarity 72.7%; Pred. No. 52;
Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 VPFATELHERL 13
   :|||||:|
Db 1 MPFATELTRKL 11

RESULT 37
US-09-987-107-20
; Sequence 20, Application US/09987107
; Patent No. US20020156007A1
; GENERAL INFORMATION:
; APPLICANT: GRAVERSEN, Jonas
; APPLICANT: MOESTRUP, Soren
; TITLE OF INVENTION: APOLIPOPROTEINS ANALOGUES
; FILE REFERENCE: GRAVERSENIA
; CURRENT APPLICATION NUMBER: US/09/987,107
; CURRENT FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/264,022
; PRIOR FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: DK PA2001 00057
; PRIOR FILING DATE: 2001-01-15
; PRIOR APPLICATION NUMBER: DK PA2000 01682
; PRIOR FILING DATE: 2000-11-10
; NUMBER OF SEQ ID NOS: 91
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 20
; LENGTH: 266
; TYPE: PRT
; ORGANISM: Oryctolagus cuniculus
US-09-987-107-20

Query Match      60.6%; Score 40; DB 9; Length 266;
Best Local Similarity 53.8%; Pred. No. 58;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 KLVPFATELHERL 13
   |||:|:|:|
Db 185 KLAPYSNELQQRL 197

RESULT 38
US-10-987-454-17
; Sequence 17, Application US/10987454
; Publication No. US20050172359A1
; GENERAL INFORMATION:
; APPLICANT: Reid, Alexandra
; APPLICANT: Moloney, Maurice
; TITLE OF INVENTION: METHODS FOR THE PRODUCTION OF APOLIPOPROTEINS IN TRANSGENIC
; TITLE OF INVENTION: PLANTS
; FILE REFERENCE: 9369-311
; CURRENT APPLICATION NUMBER: US/10/987,454
; CURRENT FILING DATE: 2004-11-15
; PRIOR APPLICATION NUMBER: US 60/519,606
; PRIOR FILING DATE: 2003-11-14
; PRIOR APPLICATION NUMBER: US 60/579,733
; PRIOR FILING DATE: 2004-06-16
; NUMBER OF SEQ ID NOS: 251
; SOFTWARE: Patentin version 3.3
; SEQ ID NO 17
; LENGTH: 266
; TYPE: PRT
```

; ORGANISM: Oryctolagus cuniculus  
US-10-987-454-17

Query Match 60.6%; Score 40; DB 18; Length 266;  
Best Local Similarity 53.8%; Pred. No. 58;  
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 KLVPFATLHERL 13  
|||::|||  
Db 185 KLAPYSNELQRL 197

## RESULT 39

US-11-017-037-20  
; Sequence 20, Application US/11017037  
; Publication No. US20050096277A1  
; GENERAL INFORMATION:  
; APPLICANT: GRAVERSEN, Jonas  
; APPLICANT: MOESTRUP, Soren  
; TITLE OF INVENTION: APOLIPOPROTEINS ANALOGUES  
; FILE REFERENCE: GRAVERSENIA  
; CURRENT APPLICATION NUMBER: US/11/017,037  
; CURRENT FILING DATE: 2004-12-21  
; PRIOR APPLICATION NUMBER: US/09/987,107  
; PRIOR FILING DATE: 2001-11-13  
; PRIOR APPLICATION NUMBER: US 60/264,022  
; PRIOR FILING DATE: 2001-01-26  
; PRIOR APPLICATION NUMBER: DK PA2001 00057  
; PRIOR FILING DATE: 2001-01-15  
; PRIOR APPLICATION NUMBER: DK PA2000 01682  
; PRIOR FILING DATE: 2000-11-10  
; NUMBER OF SEQ ID NOS: 91  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 20  
; LENGTH: 266  
; TYPE: PRT  
; ORGANISM: Oryctolagus cuniculus  
US-11-017-037-20

Query Match 60.6%; Score 40; DB 19; Length 266;  
Best Local Similarity 53.8%; Pred. No. 58;  
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 KLVPFATLHERL 13  
|||::|||  
Db 185 KLAPYSNELQRL 197

## RESULT 40

US-11-017-059-20  
; Sequence 20, Application US/11017059  
; Publication No. US20050142639A1  
; GENERAL INFORMATION:  
; APPLICANT: GRAVERSEN, Jonas  
; APPLICANT: MOESTRUP, Soren  
; TITLE OF INVENTION: APOLIPOPROTEINS ANALOGUES  
; FILE REFERENCE: GRAVERSENIA  
; CURRENT APPLICATION NUMBER: US/11/017,059  
; CURRENT FILING DATE: 2004-12-21  
; PRIOR APPLICATION NUMBER: US/09/987,107  
; PRIOR FILING DATE: 2001-11-13  
; PRIOR APPLICATION NUMBER: US 60/264,022  
; PRIOR FILING DATE: 2001-01-26  
; PRIOR APPLICATION NUMBER: DK PA2001 00057  
; PRIOR FILING DATE: 2001-01-15  
; PRIOR APPLICATION NUMBER: DK PA2000 01682  
; PRIOR FILING DATE: 2000-11-10  
; NUMBER OF SEQ ID NOS: 91  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 20  
; LENGTH: 266  
; TYPE: PRT  
; ORGANISM: Oryctolagus cuniculus

## US-11-017-059-20

Query Match 60.6%; Score 40; DB 20; Length 266;  
Best Local Similarity 53.8%; Pred. No. 58;  
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 KLVPFATLHERL 13  
|||::|||  
Db 185 KLAPYSNELQRL 197

## RESULT 41

US-10-437-963-195972  
; Sequence 195972, Application US/10437963  
; Publication No. US20040123343A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Wu, Wei  
; APPLICANT: Boukharov, Andrey A.  
; APPLICANT: Barbazuk, Brad  
; APPLICANT: Li, Ping  
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53221)B  
; CURRENT APPLICATION NUMBER: US/10/437,963  
; CURRENT FILING DATE: 2003-05-14  
; NUMBER OF SEQ ID NOS: 204966  
; SEQ ID NO 195972  
; LENGTH: 1809  
; TYPE: PRT  
; ORGANISM: Oryza sativa  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_91869C.1.pep  
US-10-437-963-195972

Query Match 60.6%; Score 40; DB 16; Length 1809;  
Best Local Similarity 53.8%; Pred. No. 5e+02;  
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 KLVPFATLHERL 13  
|||::|||  
Db 464 KLISFVDEAHERI 476

## RESULT 42

US-10-437-963-177544  
; Sequence 177544, Application US/10437963  
; Publication No. US20040123343A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Wu, Wei  
; APPLICANT: Boukharov, Andrey A.  
; APPLICANT: Barbazuk, Brad  
; APPLICANT: Li, Ping  
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53221)B  
; CURRENT APPLICATION NUMBER: US/10/437,963  
; CURRENT FILING DATE: 2003-05-14  
; NUMBER OF SEQ ID NOS: 204966  
; SEQ ID NO 177544  
; LENGTH: 86  
; TYPE: PRT  
; ORGANISM: Oryza sativa  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_75185C.1.pep  
US-10-437-963-177544

```

; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_70397C.1.pcp
US-10-425-115--336871

Query Match          59.1%; Score 39; DB 16; Length 143;
Best Local Similarity 63.6%; Pred. NO. 45;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy      2 LVPFTELHER 12
Db      133 IVRFCTEVHER 143
      :| | | | |
      :| | | | |

RESULT 45
US-09-815-242-10585
; Sequence 10585, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlssen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCES: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10585
; LENGTH: 220
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-09-815-242-10585

Query Match          59.1%; Score 39; DB 9; Length 220;
Best Local Similarity 87.5%; Pred. NO. 72;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      4 PFATELHE 11
Db      22 PFTELHE 29
      :| | | | |
      :| | | | |

Search completed: September 1, 2005, 16:45:05
Job time : 164 secs

```

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 1, 2005, 16:12:34 ; Search time 166 Seconds  
(without alignments)  
30.288 Million cell updates/sec

Title: US-09-993-366-1

Perfect score: 66

Sequence: 1 KLVPPFATLHERL 13

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database :

A Geneseq\_16Dec04:\*  
1: geneseqp1980s:\*  
2: geneseqp1990s:\*  
3: geneseqp2000s:\*  
4: geneseqp2001s:\*  
5: geneseqp2002s:\*  
6: geneseqp2003as:\*  
7: geneseqp2003bs:\*  
8: geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	66	100.0	13	6	ABR62228 Apolipop
2	66	100.0	13	8	ADE52384 Human apo
3	66	100.0	85	8	ADP87442 Human apo
4	66	100.0	85	8	ADP87439 Pig apoli
5	66	100.0	85	8	ADP87440 Baboon ap
6	66	100.0	85	8	ADP87441 Monkey ap
7	66	100.0	183	2	AAR39482 Human apo
8	66	100.0	333	2	AAR39495 Human apo
9	66	100.0	333	2	AAR39488 Human apo
10	66	100.0	333	2	AAR39481 Human apo
11	66	100.0	333	2	AAR39497 Human apo
12	66	100.0	333	2	AAR39490 Human apo
13	66	100.0	337	2	AAR39492 Human apo
14	66	100.0	337	2	AAR39494 Human apo
15	66	100.0	342	2	AAR39487 Human apo
16	66	100.0	342	2	AAR39491 Human apo
17	66	100.0	342	2	AAR39489 Human apo
18	66	100.0	342	2	AAR39498 Human apo
19	66	100.0	342	2	AAR39496 Human apo
20	66	100.0	346	2	AAR39493 Human apo
21	66	100.0	363	2	AAR39478 Human apo
22	66	100.0	363	2	AAR39479 Human apo
23	66	100.0	373	2	AAR39486 Human apo
24	66	100.0	377	2	AAR39502 Human apo
25	66	100.0	377	2	AAR39443 Human apo

26	66	100.0	377	2	AAR45244	Human apo
27	66	100.0	377	2	AAR39501	Human apo
28	66	100.0	377	2	AAR45242	Human apo
29	66	100.0	377	2	AAR39480	Human apo
30	66	100.0	377	2	AAR39499	Human apo
31	66	100.0	377	2	AAR39500	Human apo
32	66	100.0	377	2	AAR45243	Human apo
33	66	100.0	382	4	AAB90663	Human sec
34	66	100.0	396	4	AAB90664	Human sec
35	66	100.0	396	5	AAOI5885	Human apo
36	66	100.0	396	5	AAUI0862	Human apo
37	66	100.0	396	5	AAUI0863	Human apo
38	66	100.0	396	5	AAUI0865	Human apo
39	66	100.0	396	5	AAUI0867	Human apo
40	66	100.0	396	5	AAUI0866	Human apo
41	66	100.0	396	5	AAUI0869	Human apo
42	66	100.0	396	5	AAUI0861	Human apo
43	66	100.0	396	5	AAUI0868	Human apo
44	66	100.0	396	5	AAUI0860	Human apo
45	66	100.0	396	5	AAUI0864	Human apo
46	66	100.0	396	7	ADE60278	Human Pro
47	66	100.0	396	8	ADQ39968	Human myo
48	66	100.0	396	8	ADQ39969	Human myo
49	53	80.3	328	2	AAR39484	Human apo
50	53	80.3	337	2	AAR39485	Human apo
51	45	68.2	85	8	ADP87438	Chicken a
52	42	63.6	438	6	ABM69115	Photorhab
53	42	63.6	434	4	ABM66417	Drosophil
54	41	62.1	258	7	ADF30720	Rat angio
55	41	62.1	258	7	ADF30718	Rat angio
56	41	62.1	258	7	ADF30716	Rat angio
57	41	62.1	472	7	ADG33820	Actinomy
58	40	60.6	49	6	ABR44041	Rabbit Ap
59	40	60.6	242	4	ABG28667	Novel hum
60	39	59.1	93	6	ABM69659	Photorhab
61	39	59.1	220	6	AAU34992	Protein e
62	39	59.1	220	6	ABU29095	Bacterial
63	39	59.1	397	8	ADN21067	Bacterial
64	39	59.1	397	8	ADN21068	Bacterial
65	39	59.1	815	4	ABG19101	Novel hum
66	39	59.1	2796	6	ABU79133	Mycobacte
67	39	59.1	2796	7	ADP43394	Mycobacte
68	39	59.1	2796	8	ADQ26440	Mycobacte
69	39	59.1	2796	8	ADR88334	Mycobacte
70	39	59.1	3069	8	ADQ26432	Mycobacte
71	39	59.1	3069	8	ADR88327	Mycobacte
72	39	59.1	3070	4	AAG81195	Mycobacte
73	39	59.1	3076	8	ADQ26433	Mycobacte
74	39	59.1	3076	8	ADR88328	Mycobacte
75	38	57.6	15	2	AAV42559	Apolipop
76	38	57.6	213	4	ABG03485	Novel hum
77	38	57.6	243	5	ADF04137	Bacterial
78	38	57.6	261	5	AAU99616	Agrobacte
79	38	57.6	262	7	ADJ83088	Apolipop
80	38	57.6	314	7	ADC23743	Protein s
81	38	57.6	314	8	ADH35844	Chemical
82	38	57.6	314	8	ADG93545	Nitrilase
83	38	57.6	314	8	ADI62142	Nitrilase
84	38	57.6	314	8	ADI62142	Nitrilase
85	38	57.6	319	7	ADM25798	Hyperther
86	38	57.6	391	7	ADE60276	Rat Prote
87	38	57.6	391	7	ADF30675	Rat angio
88	38	57.6	391	7	ADF30677	Rat angio
89	38	57.6	391	7	ADF30679	Rat angio
90	38	57.6	405	8	ADS25206	Bacterial
91	38	57.6	405	8	ADS44961	Bacterial
92	38	57.6	405	8	ADS25821	Bacterial
93	38	57.6	405	8	ADS26010	Bacterial
94	38	57.6	405	8	ADS22551	Bacterial
95	38	57.6	428	6	ABU17846	Protein e
96	38	57.6	428	6	ADS28450	Bacterial
97	38	57.6	429	6	ABU17531	Protein e
98	38	57.6	473	8	ADJ49354	Oil-assoc

99 38 57.6 492 7 ABO67691 Abo67691 Klebsiell  
100 38 57.6 571 7 ADC51490 Bacterial

## ALIGNMENTS

```

RESULT 1
ABR62228
ID ABR62228 standard; peptide; 13 AA.
XX AC
XX AC ABR62228;
XX 08-SEP-2003 (first entry)
XX Apolipoprotein marker predictive of insulin resistance.
XX DE
XX Apolipoprotein; human; insulin resistance; marker.
XX KW
XX Homo sapiens.
XX OS
XX WO2003046000-A2.
XX PN
XX 05-JUN-2003.
XX PD
XX 31-OCT-2002; 2002WO-CA001660.
XX PF
XX 21-NOV-2001; 2001US-00993366.
XX PR
XX (SYNX-) SYN.X PHARMA INC.
XX PA
XX Jackowski G, Marshall J;
XX PI
XX WPI; 2003-493400/46.
XX DR
XX New biopolymer marker, useful for indicating, for determining risk-
PT assessment or for identifying therapeutic avenues related to, a disease
- PT state e.g., Alzheimer's disease.
XX XX
XX Claim 1; Page 43; 43pp; English.
XX PS
XX The present invention relates to the use of mass spectrometry to
CC elucidate particular biopolymer markers indicative or predictive of a
CC particular disease state, and especially biopolymer markers whose up-
CC regulation, down-regulation or relative presence in disease versus normal
CC states has been determined to be useful in disease state assessment and
CC therapeutic target recognition, development and validation. As a result
CC of these procedures, the present apolipoprotein A-IV precursor protein
CC was identified as a marker predictive of insulin resistance. The presence
CC or absence of such markers can be used in disease risk assessment and in
CC the development of therapeutic avenues against the disease
XX SQ
Query Match 100.0%; Score 66; DB 6; Length 13;
Best Local Similarity 100.0%; Pred. No. 2e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLVPFFATELHERL 13
Db 1 KLVPFFATELHERL 13

RESULT 2
ADE52384
ID ADE52384 standard; peptide; 13 AA.
XX AC
XX ADE52384;
XX AC
XX 29-JAN-2004 (first entry)
XX DT
XX Human apolipoprotein biopolymer marker #2.
XX DE
XX
XX

```

```

KW human; apolipoprotein biopolymer marker; disease state regulation;
KW therapeutic avenue.
XX OS
XX Homo sapiens.
XX XX
XX US2003100014-A1.
XX XX
XX 29-MAY-2003.
XX XX
XX 23-NOV-2001; 2001US-00991809.
XX PF
XX 23-NOV-2001; 2001US-00991809.
XX PR
XX (JACK/) JACKOMSKI G.
XX PA (MARS/) MARSHALL J.
XX XX
XX Jackowski G, Marshall J;
XX PI
XX WPI; 2004-031198/03.
XX DR
XX New apolipoprotein biopolymer marker useful for indicating at least one
XX particular disease state, such as, Type II diabetes, using e.g. mass
XX spectrometric analysis.
XX PS
XX Claim 1; Page 11; 17pp; English.
XX XX
XX The invention relates to an apolipoprotein biopolymer marker or its
CC analyte useful in indicating a particular disease state. The
CC apolipoprotein biopolymer marker is useful for regulating a disease state
CC by controlling the presence or absence of apolipoprotein biopolymer
CC marker. The apolipoprotein biopolymer marker or a diagnostic kit is
CC useful for identifying therapeutic avenues related to a disease state,
CC which involves conducting an analysis by using the diagnostic kit and
CC interacting with apolipoprotein biopolymer marker. The present sequence
CC represents the amino acid sequence of a human apolipoprotein biopolymer
CC marker.
XX CC
XX Sequence 13 AA;
XX SQ
Query Match 100.0%; Score 66; DB 8; Length 13;
Best Local Similarity 100.0%; Pred. No. 2e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLVPFFATELHERL 13
Db 1 KLVPFFATELHERL 13

RESULT 3
ADP87442
ID ADP87442 standard; protein; 85 AA.
XX AC
XX ADP87442;
XX AC
XX 09-SEP-2004 (first entry)
XX DT
XX Human apolipoprotein A (APO-A) precursor protein.
XX DE
XX Human apolipoprotein A (APO-A) precursor protein.
XX XX
XX Browseable database system; ontology; protein analysis;
XX KW gene product classification; genomic analysis; apolipoprotein A; APO-A;
XX XX human.
XX OS
XX Homo sapiens.
XX OS
XX WO2004053769-A2.
XX PN
XX 24-JUN-2004.
XX PD
XX 09-DEC-2003; 2003WO-US038935.
XX PF
XX 09-DEC-2002; 2002US-0431879P.
XX PR
XX (APPL-) APPLERA CORP.
XX XX
XX

```



XX Thomas PD, Kejariwal A, Campbell MJ, Mi H, Diemer K, Guo N;  
PI Ladunga I, Lazareva B, Muruganujan A, Rabkin S, Vandergriff J;  
PI Doremieux O;  
XX WPI; 2004-480967/45.  
XX  
XX Browseable database system, for e.g. analyzing protein sequences, and  
PT predicting a biological role, comprises a database, an ontology of  
PT categories of biological functions, an input receptive, a recognizer, and  
PT an output.  
XX  
XX Disclosure; SEQ ID NO 49; 113pp; English.  
XX  
XX The invention relates to a browseable database system for use with  
CC biological information, comprises a database, an ontology of categories  
CC of biological functions, an input receptive, a recognizer and an output.  
CC The invention also provides a method of operation for use with a  
CC browseable biological database system. The browseable database system and  
CC methods are useful in analysing protein sequences, classifying gene  
CC products, predicting biological role for pathway building, enhancing  
CC interpretation of expression information, providing protein function and  
CC in facilitating comparative genomic analysis. The present sequence is  
CC human apolipoprotein A (APO-A) IV precursor protein. This sequence is  
CC used to illustrate the method of the invention.  
XX  
XX SQ Sequence 85 AA;  
  
XX Query Match 100.0%; Score 66; DB 8; Length 85;  
XX Best Local Similarity 100.0%; Pred. No. 0.00017;  
XX Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
XX Qy 1 KLVPFATLHERL 13  
XX |  
XX 28 KLVPFATLHERL 40  
  
XX  
XX RESULT 4  
XX ADP87439  
XX ID ADP87439 standard; protein; 85 AA.  
XX AC ADP87439;  
XX DT 09-SEP-2004 (first entry)  
XX DE Pig apolipoprotein A (APO-A) IV precursor protein.  
XX Browseable database system; ontology; protein analysis;  
XX gene product classification; genomic analysis; apolipoprotein A; APO-A;  
XX Pig.  
XX OS Sus scrofa.  
XX PI Ladunga I, Lazareva B, Muruganujan A, Rabkin S, Vandergriff J;  
XX PI Doremieux O;  
XX WPI; 2004-480967/45.  
XX  
XX Browseable database system, for e.g. analyzing protein sequences, and  
PT predicting a biological role, comprises a database, an ontology of  
PT categories of biological functions, an input receptive, a recognizer, and  
PT an output.

XX Disclosure; SEQ ID NO 46; 113pp; English.  
XX  
XX The invention relates to a browseable database system for use with  
CC biological information, comprises a database, an ontology of categories  
CC of biological functions, an input receptive, a recognizer and an output.  
CC The invention also provides a method of operation for use with a  
CC browseable biological database system. The browseable database system and  
CC methods are useful in analysing protein sequences, classifying gene  
CC products, predicting biological role for pathway building, enhancing  
CC interpretation of expression information, providing protein function and  
CC in facilitating comparative genomic analysis. The present sequence is pig  
CC apolipoprotein A (APO-A) IV precursor protein. This sequence is used to  
CC illustrate the method of the invention.  
XX  
XX SQ Sequence 85 AA;  
  
XX Query Match 100.0%; Score 66; DB 8; Length 85;  
XX Best Local Similarity 100.0%; Pred. No. 0.00017;  
XX Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
XX Qy 1 KLVPFATLHERL 13  
XX |  
XX 28 KLVPFATLHERL 40  
  
XX  
XX RESULT 5  
XX ADP87440  
XX ID ADP87440 standard; protein; 85 AA.  
XX AC ADP87440;  
XX DT 09-SEP-2004 (first entry)  
XX DE Baboon apolipoprotein A (APO-A) IV precursor protein.  
XX Browseable database system; ontology; protein analysis;  
XX gene product classification; genomic analysis; apolipoprotein A; APO-A;  
XX baboon.  
XX OS Papio anubis.  
XX PI Ladunga I, Lazareva B, Muruganujan A, Rabkin S, Vandergriff J;  
XX PI Doremieux O;  
XX WPI; 2004-480967/45.  
XX  
XX Browseable database system, for e.g. analyzing protein sequences, and  
PT predicting a biological role, comprises a database, an ontology of  
PT categories of biological functions, an input receptive, a recognizer, and  
PT an output.  
XX  
XX Disclosure; SEQ ID NO 47; 113pp; English.  
XX  
XX The invention relates to a browseable database system for use with  
CC biological information, comprises a database, an ontology of categories  
CC of biological functions, an input receptive, a recognizer and an output.  
CC The invention also provides a method of operation for use with a  
CC browseable biological database system. The browseable database system and  
CC methods are useful in analysing protein sequences, classifying gene  
CC products, predicting biological role for pathway building, enhancing  
CC interpretation of expression information, providing protein function and  
CC in facilitating comparative genomic analysis. The present sequence is  
CC human apolipoprotein A (APO-A) IV precursor protein. This sequence is  
CC used to illustrate the method of the invention.



XX AAR39495;  
 AC 25-MAR-2003 (revised)  
 DT 28-JAN-1994 (first entry)  
 XX  
 DE Human apoAIV muten P(delta h13-14).  
 XX  
 KW apolipoprotein AIV; apoAIV; hypercholesterolaemia; atherosclerosis;  
 KW coronary disease; chylomicron; cholesterol transport; plaque deposition;  
 KW lecithin-cholesterol-acyltransferase; LCAT.  
 XX  
 OS Synthetic.  
 XX  
 PH Key Location/Qualifiers  
 FT 14..40  
 FT /note= "helix 1"  
 FT 41..62  
 FT /note= "helix 2"  
 FT 63..95  
 FT /note= "helix 3"  
 FT 96..117  
 FT /note= "helix 4"  
 FT 118..139  
 FT /note= "helix 5"  
 FT 140..161  
 FT /note= "helix 6"  
 FT 162..183  
 FT /note= "helix 7"  
 FT 184..205  
 FT /note= "helix 8"  
 FT 206..227  
 FT /note= "helix 9"  
 FT 228..249  
 FT /note= "helix 10"  
 FT 250..267  
 FT /note= "helix 11"  
 FT 268..289  
 FT /note= "helix 12"  
 FT Misc-difference 289..290  
 FT /note= "helices 13-14 have been deleted from between  
 FT these two sites"  
 FT 290..333  
 FT /note= "helix 15"  
 XX  
 PN WO9315198-A1.  
 XX  
 PD 05-AUG-1993.  
 XX  
 PF 26-JAN-1993; 93WO-FR0000073.  
 XX  
 PR 27-JAN-1992; 92FR-00000806.  
 XX  
 PA (RHON ) RHONE POULENC RORER SA.  
 XX  
 PI Deneffe P, Guinet F, Latta M, Murry-Brelier A;  
 XX  
 XX WPI; 1993-258676/32.  
 DR  
 XX New apo:lipoprotein AIV mutant polypeptide(s) - useful for drug design  
 XX and/or treatment of hypercholesterolaemia.  
 XX  
 PS Claim 7; Page 31-32 and Page 5; 42pp; French.  
 XX  
 CC Human apolipoprotein AIV is a major component of chylomicrons in lymph  
 CC but is mainly found in the plasma in unassociated form. The apoAIV  
 CC mediates inverse transport of cholesterol. The invention covers  
 CC polypeptides which are derived from apoAIV by deletion of at least 10  
 CC terminal amino acids, by deletion of a helix or pair of helices, by  
 CC addition of a heterologous polypeptide portion or by a point mutation.  
 CC Such mutants are useful in the design of hypocholesterolaemic drugs to  
 CC treat hypercholesterolaemia and atherosclerosis. AAR39495 is a  
 CC specifically claimed muten and the sequence has been compiled from the

CC wild-type sequence (see AAR39443) and the description given in the  
 CC disclosure; the mutant sequence is not shown in the specification.  
 CC (Updated on 25-MAR-2003 to correct PN field.)  
 XX  
 SQ Sequence 333 AA;  
 Query Match 100.0%; Score 66; DB 2; Length 333;  
 Best Local Similarity 100.0%; Pred. No. 0.00078;  
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 KLVFFATELHERL 13  
 DB 60 KLVFFATELHERL 72  
 RESULT 9  
 ID AAR39488  
 AC AAR39488;  
 XX  
 XX 25-MAR-2003 (revised)  
 DT 28-JAN-1994 (first entry)  
 XX  
 DE Human apoAIV muten P(delta h7-8).  
 XX  
 KW apolipoprotein AIV; apoAIV; hypercholesterolaemia; atherosclerosis;  
 KW coronary disease; chylomicron; cholesterol transport; plaque deposition;  
 KW lecithin-cholesterol-acyltransferase; LCAT.  
 XX  
 OS Synthetic.  
 XX  
 PH Key Location/Qualifiers  
 FT 14..40  
 FT /note= "helix 1"  
 FT 41..62  
 FT /note= "helix 2"  
 FT 63..95  
 FT /note= "helix 3"  
 FT 96..117  
 FT /note= "helix 4"  
 FT 118..139  
 FT /note= "helix 5"  
 FT 140..161  
 FT /note= "helix 6"  
 FT Misc-difference 161..162  
 FT /note= "helices 7-8 have been deleted from between these  
 FT two sites"  
 FT 162..183  
 FT /note= "helix 9"  
 FT 184..205  
 FT /note= "helix 10"  
 FT 206..223  
 FT /note= "helix 11"  
 FT 224..245  
 FT /note= "helix 12"  
 FT 246..267  
 FT /note= "helix 13"  
 FT 268..289  
 FT /note= "helix 14"  
 FT 290..333  
 FT /note= "helix 15"  
 XX  
 PN WO9315198-A1.  
 XX  
 PD 05-AUG-1993.  
 XX  
 PF 26-JAN-1993; 93WO-FR0000073.  
 XX  
 PR 27-JAN-1992; 92FR-00000806.  
 XX  
 PA (RHON ) RHONE POULENC RORER SA.  
 XX  
 PI Deneffe P, Guinet F, Latta M, Murry-Brelier A;  
 XX  
 XX WPI; 1993-258676/32.  
 DR  
 XX New apo:lipoprotein AIV mutant polypeptide(s) - useful for drug design  
 XX and/or treatment of hypercholesterolaemia.  
 XX  
 PS Claim 7; Page 31-32 and Page 5; 42pp; French.  
 XX  
 CC Human apolipoprotein AIV is a major component of chylomicrons in lymph  
 CC but is mainly found in the plasma in unassociated form. The apoAIV  
 CC mediates inverse transport of cholesterol. The invention covers  
 CC polypeptides which are derived from apoAIV by deletion of at least 10  
 CC terminal amino acids, by deletion of a helix or pair of helices, by  
 CC addition of a heterologous polypeptide portion or by a point mutation.  
 CC Such mutants are useful in the design of hypocholesterolaemic drugs to  
 CC treat hypercholesterolaemia and atherosclerosis. AAR39495 is a  
 CC specifically claimed muten and the sequence has been compiled from the

PI Denefle P, Guinet F, Latta M, Murry-Brelrier A;  
XX WPI; 1993-258676/32.  
DR New apo:lipoprotein AIV mutant polypeptide(s) - useful for drug design  
XX and/or treatment of hypercholesterolaemia.  
XX  
XX Claim 7; Page 31-32 and Page 4; 42pp; French.  
XX  
XX Human apolipoprotein AIV is a major component of chylomicrons in lymph  
CC but is mainly found in the plasma in unassociated form. The apoAIV  
CC mediates inverse transport of cholesterol. The invention covers  
CC polypeptides which are derived from apoAIV by deletion of at least 10  
CC terminal amino acids, by deletion of a helix or pair of helices, by  
CC addition of a heterologous polypeptide portion or by a point mutation.  
CC Such muteins are useful in the design of hypocholesterolaemic drugs to  
CC treat hypercholesterolaemia and atherosclerosis. AAR39481 is a  
CC specifically claimed mutein and the sequence has been compiled from the  
CC wild-type sequence (see AAR39443) and the description given in the  
CC disclosure; the mutant sequence is not shown in the specification.  
CC (Updated on 25-MAR-2003 to correct PN field.)  
XX  
XX SQ Sequence 333 AA;

Query Match 100.0%; Score 66; DB 2; Length 333;  
Best Local Similarity 100.0%; Pred. No. 0.00078;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KLVPFATELHERL 13  
Db 60 KLVPFATELHERL 72

RESULT 10  
AAR39481  
ID AAR39481 standard; protein; 333 AA.  
XX  
XX AAR39481;  
XX  
XX 25-MAR-2003 (revised)  
DT 28-JAN-1994 (first entry)  
DE Human apoAIV mutein P(deltaC44).  
XX  
XX apolipoprotein AIV; apoAIV; hypercholesterolaemia; atherosclerosis;  
KW coronary disease; chylomicron; cholesterol transport; plaque deposition;  
KW lecithin-cholesterol-acyltransferase; LCAT.  
XX  
XX Synthetic.

Key Location/Qualifiers  
FT Protein 2..333  
FT /label= P(deltaC44)  
FT /note= "the 44 C-terminal amino acids of human apoAIV  
FT have been deleted"

WO9315198-A1.  
XX  
XX 05-AUG-1993.  
XX  
XX 26-JAN-1993; 93WO-FR000073.  
XX  
XX 27-JAN-1992; 92FR-00000806.  
XX  
XX (RHON ) RHONE POULENC RORER SA.  
XX  
XX Denefle P, Guinet F, Latta M, Murry-Brelrier A;  
XX WPI; 1993-258676/32.

XX New apo:lipoprotein AIV mutant polypeptide(s) - useful for drug design  
XX and/or treatment of hypercholesterolaemia.  
XX

PS Claim 7; Page 31-32 and Page 4; 42pp; French.  
XX  
XX Human apolipoprotein AIV is a major component of chylomicrons in lymph  
CC but is mainly found in the plasma in unassociated form. The apoAIV  
CC mediates inverse transport of cholesterol. The invention covers  
CC polypeptides which are derived from apoAIV by deletion of at least 10  
CC terminal amino acids, by deletion of a helix or pair of helices, by  
CC addition of a heterologous polypeptide portion or by a point mutation.  
CC Such muteins are useful in the design of hypocholesterolaemic drugs to  
CC treat hypercholesterolaemia and atherosclerosis. AAR39481 is a  
CC specifically claimed mutein and the sequence has been compiled from the  
CC wild-type sequence (see AAR39443) and the description given in the  
CC disclosure; the mutant sequence is not shown in the specification.  
CC (Updated on 25-MAR-2003 to correct PN field.)  
XX  
XX SQ Sequence 333 AA;

Query Match 100.0%; Score 66; DB 2; Length 333;  
Best Local Similarity 100.0%; Pred. No. 0.00078;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KLVPFATELHERL 13  
Db 60 KLVPFATELHERL 72

RESULT 11  
AAR39497  
ID AAR39497 standard; protein; 333 AA.  
XX  
XX AAR39497;  
XX

XX 25-MAR-2003 (revised)  
DT 28-JAN-1994 (first entry)  
DE Human apoAIV mutein P(delta h5-6).  
XX

XX apolipoprotein AIV; apoAIV; hypercholesterolaemia; atherosclerosis;  
KW coronary disease; chylomicron; cholesterol transport; plaque deposition;  
KW lecithin-cholesterol-acyltransferase; LCAT.  
XX  
XX Synthetic.

Key Location/Qualifiers  
FT Region 14..40  
FT /note= "helix 1"  
FT Region 41..62  
FT /note= "helix 2"  
FT Region 63..95  
FT /note= "helix 3"  
FT Region 96..117  
FT /note= "helix 4"  
FT Misc-difference 117..118  
FT /note= "helices 5-6 have been deleted from between these  
FT two sites"  
FT Region 118..139  
FT /note= "helix 7"  
FT Region 140..161  
FT /note= "helix 8"  
FT Region 162..183  
FT /note= "helix 9"  
FT Region 184..205  
FT /note= "helix 10"  
FT Region 206..223  
FT /note= "helix 11"  
FT Region 224..245  
FT /note= "helix 12"  
FT Region 246..267  
FT /note= "helix 13"  
FT Region 268..289  
FT /note= "helix 14"  
FT Region 290..333  
FT /note= "helix 15"

```

XX WO9315198-A1.
XX
XX PD
XX
XX PD
XX
XX PF 26-JAN-1993; 93WO-FR0000073.
XX PR 27-JAN-1992; 92FR-00000806.
XX
XX PA (RHON ) RHONE POULENC RORER SA.
XX
XX PI Denefle P, Guinet F, Latta M, Murry-Brelrier A;
XX WPI; 1993-258676/32.
XX
XX DR
XX PT New apo:lipoprotein AIV mutant polypeptide(s) - useful for drug design
XX PT and/or treatment of hypercholesterolaemia.
XX
XX PS Claim 7; Page 31-32 and Page 5; 42pp; French.
XX
XX CC Human apolipoprotein AIV is a major component of chylomicrons in lymph
XX CC but is mainly found in the plasma in unassociated form. The apoAIV
XX CC mediates inverse transport of cholesterol. The invention covers
XX CC polypeptides which are derived from apoAIV by deletion of at least 10
XX CC terminal amino acids, by deletion of a helix or pair of helices, by
XX CC addition of a heterologous polypeptide portion or by a point mutation.
XX CC Such muteins are useful in the design of hypocholesterolaemic drugs to
XX CC treat hypercholesterolaemia and atherosclerosis. AAR39490 is a
XX CC specifically claimed mutein and the sequence has been compiled from the
XX CC wild-type sequence (see AAR39443) and the description given in the
XX CC disclosure; the mutant sequence is not shown in the specification.
XX CC (Updated on 25-MAR-2003 to correct PN field.)
XX
XX SQ Sequence 333 AA;
XX
XX Query Match 100.0%; Score 66; DB 2; Length 333;
XX Best Local Similarity 100.0%; Pred. No. 0.00078;
XX Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX Qy 1 KLVPFATLHERL 13
XX |||||
XX Db 60 KLVPFATLHERL 72
XX
XX RESULT 12
XX ID AAR39490 standard; protein; 333 AA.
XX
XX AC AAR39490;
XX
XX DT 25-MAR-2003 (revised)
XX DT 28-JAN-1994 (first entry)
XX
XX DE Human apoAIV mutein P(delta h9-10).
XX
XX KW apolipoprotein AIV; apoAIV; hypercholesterolaemia; atherosclerosis;
XX KW coronary disease; chylomicron; cholesterol transport; plaque deposition;
XX KW lecithin-cholesterol-acyltransferase; LCAT.
XX
XX OS Synthetic.
XX
XX FH Key Location/Qualifiers
XX FT Region 14..40
XX FT /note= "helix 1"
XX FT Region 41..62
XX FT /note= "helix 2"
XX FT Region 63..95
XX FT /note= "helix 3"
XX FT Region 96..117
XX FT /note= "helix 4"
XX FT Region 118..139
XX FT /note= "helix 5"
XX FT Region 140..161

```

```

FT Region /note= "helix 6"
FT 162..183
FT /note= "helix 7"
FT 184..205
FT /note= "helix 8"
FT Misc-difference 205..206
FT /note= "helices 9-10 have been deleted from between these
FT two sites"
FT 206..223
FT /note= "helix 11"
FT 224..245
FT /note= "helix 12"
FT 246..267
FT /note= "helix 13"
FT 268..289
FT /note= "helix 14"
FT 290..333
FT /note= "helix 15"
XX
XX PN WO9315198-A1.
XX
XX PD 05-AUG-1993.
XX
XX PF 26-JAN-1993; 93WO-FR0000073.
XX
XX PR 27-JAN-1992; 92FR-00000806.
XX
XX PA (RHON ) RHONE POULENC RORER SA.
XX
XX PI Denefle P, Guinet F, Latta M, Murry-Brelrier A;
XX WPI; 1993-258676/32.
XX
XX PT New apo:lipoprotein AIV mutant polypeptide(s) - useful for drug design
XX PT and/or treatment of hypercholesterolaemia.
XX
XX PS Claim 7; Page 31-32 and Page 5; 42pp; French.
XX
XX CC Human apolipoprotein AIV is a major component of chylomicrons in lymph
XX CC but is mainly found in the plasma in unassociated form. The apoAIV
XX CC mediates inverse transport of cholesterol. The invention covers
XX CC polypeptides which are derived from apoAIV by deletion of at least 10
XX CC terminal amino acids, by deletion of a helix or pair of helices, by
XX CC addition of a heterologous polypeptide portion or by a point mutation.
XX CC Such muteins are useful in the design of hypocholesterolaemic drugs to
XX CC treat hypercholesterolaemia and atherosclerosis. AAR39490 is a
XX CC specifically claimed mutein and the sequence has been compiled from the
XX CC wild-type sequence (see AAR39443) and the description given in the
XX CC disclosure; the mutant sequence is not shown in the specification.
XX CC (Updated on 25-MAR-2003 to correct PN field.)
XX
XX SQ Sequence 333 AA;
XX
XX Query Match 100.0%; Score 66; DB 2; Length 333;
XX Best Local Similarity 100.0%; Pred. No. 0.00078;
XX Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX Qy 1 KLVPFATLHERL 13
XX |||||
XX Db 60 KLVPFATLHERL 72
XX
XX RESULT 13
XX AAR39492
XX ID AAR39492 standard; protein; 337 AA.
XX
XX AC AAR39492;
XX
XX DT 25-MAR-2003 (revised)
XX DT 28-JAN-1994 (first entry)
XX
XX DE Human apoAIV mutein P(delta h11-12).

```

KW apolipoprotein AIV; apoAIV; hypercholesterolaemia; atherosclerosis;  
 KW coronary disease; chylomicron; cholesterol transport; plaque deposition;  
 KW lecithin-cholesterol-acyltransferase; LCAT.  
 XX  
 OS Synthetic.

XX Key Location/Qualifiers  
 FH Region 14..40  
 FT /note= "helix 1"  
 FT Region 41..62  
 FT /note= "helix 2"  
 FT Region 63..95  
 FT /note= "helix 3"  
 FT Region 96..117  
 FT /note= "helix 4"  
 FT Region 118..139  
 FT /note= "helix 5"  
 FT Region 140..161  
 FT /note= "helix 6"  
 FT Region 162..183  
 FT /note= "helix 7"  
 FT Region 184..205  
 FT /note= "helix 8"  
 FT Region 206..227  
 FT /note= "helix 9"  
 FT Region 228..249  
 FT /note= "helix 10"  
 FT Misc-difference 249..250  
 FT /note= "helices 11-12 have been deleted from between  
 FT these two sites"  
 FT Region 250..271  
 FT /note= "helix 13"  
 FT Region 272..293  
 FT /note= "helix 14"  
 FT Region 294..337  
 FT /note= "helix 15"

XX WO9315198-A1.  
 XX  
 XX 05-AUG-1993.  
 XX  
 XX 26-JAN-1993; 93WO-FR000073.  
 XX  
 XX 27-JAN-1992; 92FR-00000806.  
 XX  
 XX (RHON ) RHONE POULENC RORER SA.  
 XX  
 XX Denefle P, Guinet F, Latta M, Murry-Brelrier A;  
 XX WPI; 1993-258676/32.  
 XX  
 XX New apo:lipoprotein AIV mutant polypeptide(s) - useful for drug design  
 XX and/or treatment of hypercholesterolaemia.

XX Claim 7; Page 31-32 and Page 5; 42pp; French.  
 XX  
 XX Human apolipoprotein AIV is a major component of chylomicrons in lymph  
 XX but is mainly found in the plasma in unassociated form. The apoAIV  
 XX mediates inverse transport of cholesterol. The invention covers  
 XX polypeptides which are derived from apoAIV by deletion of at least 10  
 XX terminal amino acids, by deletion of a helix or pair of helices, by  
 XX addition of a heterologous polypeptide portion or by a point mutation.  
 XX Such mutants are useful in the design of hypocholesterolaemic drugs to  
 XX treat hypercholesterolaemia and atherosclerosis. AAR39492 is a  
 XX specifically claimed mutein and the sequence has been compiled from the  
 XX wild-type sequence (see AAR39443) and the description given in the  
 XX disclosure; the mutant sequence is not shown in the specification.  
 XX (Updated on 25-MAR-2003 to correct PN field.)

SQ Sequence 337 AA;

Query Match 100.0%; Score 66; DB 2; Length 337;  
 Best Local Similarity 100.0%; Pred. No. 0.00079;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 KLVFPATELHERL 13  
 Db 60 KLVFPATELHERL 72  
 RESULT 14  
 AAR39494  
 ID AAR39494 standard; protein; 337 AA.  
 XX  
 AC AAR39494;  
 XX  
 DT 25-MAR-2003 (revised)  
 DT 28-JAN-1994 (first entry)  
 XX  
 DE Human apoAIV mutein P(delta h11-12, L87M).  
 XX  
 KW apolipoprotein AIV; apoAIV; hypercholesterolaemia; atherosclerosis;  
 KW coronary disease; chylomicron; cholesterol transport; plaque deposition;  
 KW lecithin-cholesterol-acyltransferase; LCAT.  
 XX  
 OS Synthetic.  
 XX  
 FH Key Location/Qualifiers  
 FT Protein 2..337  
 FT /label= (delta h11-12, L87M)  
 FT /note= "helices 11-12 have been deleted from mature human  
 FT apoAIV and wild-type Leu87 has been substd. by Met"  
 FT Region 14..40  
 FT /note= "helix 1"  
 FT Region 41..62  
 FT /note= "helix 2"  
 FT Region 63..95  
 FT /note= "helix 3"  
 FT Misc-difference 87  
 FT /note= "L87M"  
 FT Region 96..117  
 FT /note= "helix 4"  
 FT Region 118..139  
 FT /note= "helix 5"  
 FT Region 140..161  
 FT /note= "helix 6"  
 FT Region 162..183  
 FT /note= "helix 7"  
 FT Region 184..205  
 FT /note= "helix 8"  
 FT Region 206..227  
 FT /note= "helix 9"  
 FT Region 228..249  
 FT /note= "helix 10"  
 FT Misc-difference 249..250  
 FT /note= "helices 11-12 have been deleted from between  
 FT these two sites"  
 FT Region 250..271  
 FT /note= "helix 13"  
 FT Region 272..293  
 FT /note= "helix 14"  
 FT Region 294..337  
 FT /note= "helix 15"  
 XX WO9315198-A1.  
 XX  
 XX 05-AUG-1993.  
 XX  
 XX 26-JAN-1993; 93WO-FR000073.  
 XX  
 XX 27-JAN-1992; 92FR-00000806.  
 XX  
 XX (RHON ) RHONE POULENC RORER SA.  
 XX  
 XX Denefle P, Guinet F, Latta M, Murry-Brelrier A;  
 XX WPI; 1993-258676/32.  
 XX  
 XX New apo:lipoprotein AIV mutant polypeptide(s) - useful for drug design  
 XX and/or treatment of hypercholesterolaemia.

DR WPI; 1993-258676/32.  
 XX  
 PT New apolipoprotein AIV mutant polypeptide(s) - useful for drug design  
 PT and/or treatment of hypercholesterolaemia.  
 XX  
 PS  
 PS Claim 7; Page 31-32 and Page 5; 42pp; French.  
 XX  
 XX Human apolipoprotein AIV is a major component of chylomicrons in lymph  
 CC but is mainly found in the plasma in unassociated form. The apoAIV  
 CC mediates inverse transport of cholesterol. The invention covers  
 CC polypeptides which are derived from apoAIV by deletion of at least 10  
 CC terminal amino acids, by deletion of a helix or pair of helices, by  
 CC addition of a heterologous polypeptide portion or by a point mutation.  
 CC Such muteins are useful in the design of hypocholesterolaemic drugs to  
 CC treat hypercholesterolaemia and atherosclerosis. AAR39494 is a  
 CC specifically claimed mutein and the sequence has been compiled from the  
 CC wild-type sequence (see AAR39443) and the description given in the  
 CC disclosure; the mutant sequence is not shown in the specification.  
 CC (Updated on 25-MAR-2003 to correct PN field.)  
 XX  
 SQ Sequence 337 AA;  
 Query Match 100.0%; Score 66; DB 2; Length 337;  
 Best Local Similarity 100.0%; Pred. No. 0.00079;  
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 KLVPPFATELHERL 13  
 DB 60 KLVPPFATELHERL 72  
 RESULT 15  
 AAR39487  
 ID AAR39487 standard; protein; 342 AA.  
 XX  
 AC AAR39487;  
 XX  
 DT 25-MAR-2003 (revised)  
 DT 28-JAN-1994 (first entry)  
 XX  
 DE Human apoAIV mutein P(tag-deltaC44).  
 KW apolipoprotein AIV; apoAIV; hypercholesterolaemia; atherosclerosis;  
 KW coronary disease; chylomicron; cholesterol transport; plaque deposition;  
 KW lecithin-cholesterol-acyltransferase; LCAT; tag purification marker.  
 XX  
 OS Synthetic.  
 Key Location/Qualifiers  
 FH Peptide 1..10  
 FT /label= tag decapeptide  
 FT /note= "allows rapid, single stage purification without  
 FT affecting the properties of the mature mutein"  
 FT 11..342  
 FT /label= P(deltaC44)  
 FT /note= "the 44 C-terminal amino acids of mature human  
 FT apoAIV have been deleted"  
 XX WO9315198-A1.  
 XX  
 XX 05-AUG-1993.  
 XX  
 XX 26-JAN-1993; 93WO-FR0000073.  
 XX  
 XX 27-JAN-1992; 92FR-00000806.  
 XX  
 XX (RHON ) RHONE POULENC RORER SA.  
 XX  
 XX Deneffe P, Guinet F, Latta M, Murry-Brelrier A;  
 XX WPI; 1993-258676/32.  
 XX  
 XX New apolipoprotein AIV mutant polypeptide(s) - useful for drug design  
 PT

PT and/or treatment of hypercholesterolaemia.  
 XX  
 PS Claim 7; Page 31-32 and Page 4; 42pp; French.  
 XX  
 XX Human apolipoprotein AIV is a major component of chylomicrons in lymph  
 CC but is mainly found in the plasma in unassociated form. The apoAIV  
 CC mediates inverse transport of cholesterol. The invention covers  
 CC polypeptides which are derived from apoAIV by deletion of at least 10  
 CC terminal amino acids, by deletion of a helix or pair of helices, by  
 CC addition of a heterologous polypeptide portion or by a point mutation.  
 CC Such muteins are useful in the design of hypocholesterolaemic drugs to  
 CC treat hypercholesterolaemia and atherosclerosis. AAR39487 is a  
 CC specifically claimed mutein and the sequence has been compiled from the  
 CC wild-type sequence (see AAR39443) and the description given in the  
 CC disclosure; the mutant sequence is not shown in the specification.  
 CC (Updated on 25-MAR-2003 to correct PN field.)  
 XX  
 SQ Sequence 342 AA;  
 Query Match 100.0%; Score 66; DB 2; Length 342;  
 Best Local Similarity 100.0%; Pred. No. 0.0008;  
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 KLVPPFATELHERL 13  
 DB 69 KLVPPFATELHERL 81  
 RESULT 16  
 AAR39491  
 ID AAR39491 standard; protein; 342 AA.  
 XX  
 AC AAR39491;  
 XX  
 DT 25-MAR-2003 (revised)  
 DT 28-JAN-1994 (first entry)  
 XX  
 DE Human apoAIV mutein P(tag-delta h9-10).  
 KW apolipoprotein AIV; apoAIV; hypercholesterolaemia; atherosclerosis;  
 KW coronary disease; chylomicron; cholesterol transport; plaque deposition;  
 KW lecithin-cholesterol-acyltransferase; LCAT; tag purification marker.  
 XX  
 OS Synthetic.  
 Key Location/Qualifiers  
 FH Peptide 1..10  
 FT /label= tag decapeptide  
 FT /note= "allows rapid, single stage purification without  
 FT affecting the properties of the mature mutein"  
 FT 23..49  
 FT /note= "helix 1"  
 FT 50..71  
 FT /note= "helix 2"  
 FT 72..104  
 FT /note= "helix 3"  
 FT 105..126  
 FT /note= "helix 4"  
 FT 127..148  
 FT /note= "helix 5"  
 FT 149..170  
 FT /note= "helix 6"  
 FT 171..192  
 FT /note= "helix 7"  
 FT 193..214  
 FT /note= "helix 8"  
 FT 214..215  
 FT /note= "helices 9-10 have been deleted from between these  
 FT two sites"  
 FT 215..232  
 FT /note= "helix 11"  
 FT 233..254  
 FT /note= "helix 12"  
 FT

FT Region 255..276  
 FT /note= "helix 13"  
 FT 277..298  
 FT /note= "helix 14"  
 FT Region 299..342  
 FT /note= "helix 15"  
 XX  
 XX WO9315198-A1.  
 XX  
 XX 05-AUG-1993.  
 XX  
 XX 26-JAN-1993; 93WO-FR0000073.  
 XX  
 XX 27-JAN-1992; 92FR-00000806.  
 XX  
 XX (RHON ) RHONE POULENC RORER SA.  
 XX  
 XX Denefle P, Guinet F, Latta M, Murry-Brelrier A;  
 XX WPI; 1993-258676/32.  
 XX  
 XX New apo-lipoprotein AIV mutant polypeptide(s) - useful for drug design  
 FT and/or treatment of hypercholesterolaemia.  
 XX  
 XX Claim 7; Page 31-32 and Page 5; 42pp; French.  
 XX  
 XX Human apolipoprotein AIV is a major component of chylomicrons in lymph  
 CC but is mainly found in the plasma in unassociated form. The apoAIV  
 CC mediates inverse transport of cholesterol. The invention covers  
 CC polypeptides which are derived from apoAIV by deletion of at least 10  
 CC terminal amino acids, by deletion of a helix or pair of helices, by  
 CC addition of a heterologous polypeptide portion or by a point mutation.  
 CC Such muteins are useful in the design of hypocholesterolaemic drugs to  
 CC treat hypercholesterolaemia and atherosclerosis. AAR39491 is a  
 CC specifically claimed mutein and the sequence has been compiled from the  
 CC wild-type sequence (see AAR39443) and the description given in the  
 CC disclosure; the mutant sequence is not shown in the specification.  
 CC (Updated on 25-MAR-2003 to correct PN field.)  
 XX  
 XX SQ Sequence 342 AA;  
 Query Match 100.0%; Score 66; DB 2; Length 342;  
 Best Local Similarity 100.0%; Pred. No. 0.0008;  
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 KLVPFATLHERL 13  
 Db 69 KLVPFATLHERL 81  
 RESULT 17  
 AAR39489  
 ID AAR39489 standard; protein; 342 AA.  
 XX  
 XX AAR39489;  
 XX  
 XX 25-MAR-2003 (revised)  
 DT 28-JAN-1994 (first entry)  
 XX  
 XX Human apoAIV mutein P(tag-delta h7-8).  
 XX  
 XX apolipoprotein AIV; apoAIV; hypercholesterolaemia; atherosclerosis;  
 KW coronary disease; chylomicron; cholesterol transport; plaque deposition;  
 KW lecithin-cholesterol-acyltransferase; LCAT; tag purification marker.  
 XX  
 XX Synthetic.  
 XX  
 XX Key Location/Qualifiers  
 FT 1..10  
 FT Peptide /label= tag decapeptide  
 FT /note= "allows rapid, single stage purification without  
 FT affecting the properties of the mature mutein"  
 FT 23..49

FT Region /note= "helix 1"  
 FT 50..71  
 FT /note= "helix 2"  
 FT Region 72..104  
 FT /note= "helix 3"  
 FT Region 105..126  
 FT /note= "helix 4"  
 FT Region 127..148  
 FT /note= "helix 5"  
 FT Region 149..170  
 FT /note= "helix 6"  
 FT Misc-difference 170..171  
 FT /note= "helices 7-8 have been deleted from between these  
 FT two sites"  
 FT 171..192  
 FT /note= "helix 9"  
 FT Region 193..214  
 FT /note= "helix 10"  
 FT Region 215..232  
 FT /note= "helix 11"  
 FT Region 233..254  
 FT /note= "helix 12"  
 FT Region 255..276  
 FT /note= "helix 13"  
 FT Region 277..298  
 FT /note= "helix 14"  
 FT Region 299..342  
 FT /note= "helix 15"  
 XX WO9315198-A1.  
 XX  
 XX 05-AUG-1993.  
 XX  
 XX 26-JAN-1993; 93WO-FR0000073.  
 XX  
 XX 27-JAN-1992; 92FR-00000806.  
 XX  
 XX (RHON ) RHONE POULENC RORER SA.  
 XX  
 XX Denefle P, Guinet F, Latta M, Murry-Brelrier A;  
 XX WPI; 1993-258676/32.  
 XX  
 XX New apo-lipoprotein AIV mutant polypeptide(s) - useful for drug design  
 FT and/or treatment of hypercholesterolaemia.  
 XX  
 XX Claim 7; Page 31-32 and Page 4; 42pp; French.  
 XX  
 XX Human apolipoprotein AIV is a major component of chylomicrons in lymph  
 CC but is mainly found in the plasma in unassociated form. The apoAIV  
 CC mediates inverse transport of cholesterol. The invention covers  
 CC polypeptides which are derived from apoAIV by deletion of at least 10  
 CC terminal amino acids, by deletion of a helix or pair of helices, by  
 CC addition of a heterologous polypeptide portion or by a point mutation.  
 CC Such muteins are useful in the design of hypocholesterolaemic drugs to  
 CC treat hypercholesterolaemia and atherosclerosis. AAR39489 is a  
 CC specifically claimed mutein and the sequence has been compiled from the  
 CC wild-type sequence (see AAR39443) and the description given in the  
 CC disclosure; the mutant sequence is not shown in the specification.  
 CC (Updated on 25-MAR-2003 to correct PN field.)  
 XX  
 XX SQ Sequence 342 AA;  
 Query Match 100.0%; Score 66; DB 2; Length 342;  
 Best Local Similarity 100.0%; Pred. No. 0.0008;  
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 KLVPFATLHERL 13  
 Db 69 KLVPFATLHERL 81  
 RESULT 18



AAR39498  
ID AAR39498 standard; protein; 342 AA.  
XX  
AC AAR39498;  
XX  
DT 25-MAR-2003 (revised)  
DT 28-JAN-1994 (first entry)  
XX  
DE Human apoAIV mutin P(tag-delta h5-6).  
XX  
KW apolipoprotein AIV; apoAIV; hypercholesterolaemia; atherosclerosis;  
KW coronary disease; chylomicron; cholesterol transport; plaque deposition;  
KW lecithin-cholesterol-acyltransferase; LCAT; tag purification marker.  
XX  
OS Synthetic.  
XX  
FH Key Location/Qualifiers  
FT Peptide 1..10  
FT /label= tag decapeptide  
FT /note= "allows rapid, single stage purification without  
FT affecting the properties of the mature mutin"  
FT Region 23..49  
FT /note= "helix 1"  
FT Region 50..71  
FT /note= "helix 2"  
FT Region 72..104  
FT /note= "helix 3"  
FT Region 105..126  
FT /note= "helix 4"  
FT Misc-difference 126..127  
FT /note= "helices 5-6 have been deleted from between these  
FT two sites"  
FT Region 127..148  
FT /note= "helix 7"  
FT Region 149..170  
FT /note= "helix 8"  
FT Region 171..192  
FT /note= "helix 9"  
FT Region 193..214  
FT /note= "helix 10"  
FT Region 215..232  
FT /note= "helix 11"  
FT Region 233..254  
FT /note= "helix 12"  
FT Region 255..276  
FT /note= "helix 13"  
FT Region 277..298  
FT /note= "helix 14"  
FT Region 299..342  
FT /note= "helix 15"  
XX  
PN WO9315198-A1.  
XX  
PD 05-AUG-1993.  
XX  
PF 26-JAN-1993; 93WO-FR000073.  
XX  
PR 27-JAN-1992; 92FR-0000806.  
XX  
PA (RHON ) RHONE POULENC RORER SA.  
XX  
PI Denefle P, Guinet F, Latta M, Murry-Brellier A;  
XX  
DR WPI; 1993-258676/32.  
XX  
PT New apo:lipoprotein AIV mutant polypeptide(s) - useful for drug design  
PT and/or treatment of hypercholesterolaemia.  
XX  
PS Claim 7; Page 31-32 and Page 5; 42pp; French.  
XX  
CC Human apolipoprotein AIV is a major component of chylomicrons in lymph  
CC but is mainly found in the plasma in unassociated form. The apoAIV  
CC mediates inverse transport of cholesterol. The invention covers

polypeptides which are derived from apoAIV by deletion of at least 10  
terminal amino acids, by deletion of a helix or pair of helices, by  
addition of a heterologous polypeptide portion or by a point mutation.  
Such mutins are useful in the design of hypocholesterolaemic drugs to  
treat hypercholesterolaemia and atherosclerosis. AAR39498 is a  
specifically claimed mutin and the sequence has been compiled from the  
wild-type sequence (see AAR39443) and the description given in the  
disclosure; the mutant sequence is not shown in the specification.  
(Updated on 25-MAR-2003 to correct PN field.)  
XX  
SQ Sequence 342 AA;  
Query Match 100.0%; Score 66; DB 2; Length 342;  
Best Local Similarity 100.0%; Pred. No. 0.0008;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 KLVFFATELHERL 13  
DB 69 KLVFFATELHERL 81  
RESULT 19  
AAR39496  
ID AAR39496 standard; protein; 342 AA.  
XX  
AC AAR39496;  
XX  
DT 25-MAR-2003 (revised)  
DT 28-JAN-1994 (first entry)  
XX  
DE Human apoAIV mutin P(tag-delta h13-14).  
XX  
KW apolipoprotein AIV; apoAIV; hypercholesterolaemia; atherosclerosis;  
KW coronary disease; chylomicron; cholesterol transport; plaque deposition;  
KW lecithin-cholesterol-acyltransferase; LCAT; tag purification marker.  
XX  
OS Synthetic.  
XX  
FH Key Location/Qualifiers  
FT Peptide 1..10  
FT /label= tag decapeptide  
FT /note= "allows rapid, single stage purification without  
FT affecting the properties of the mature mutin"  
FT Region 23..49  
FT /note= "helix 1"  
FT Region 50..71  
FT /note= "helix 2"  
FT Region 72..104  
FT /note= "helix 3"  
FT Region 105..126  
FT /note= "helix 4"  
FT Region 127..148  
FT /note= "helix 5"  
FT Region 149..170  
FT /note= "helix 6"  
FT Region 171..192  
FT /note= "helix 7"  
FT Region 193..214  
FT /note= "helix 8"  
FT Region 215..236  
FT /note= "helix 9"  
FT Region 237..258  
FT /note= "helix 10"  
FT Region 259..276  
FT /note= "helix 11"  
FT Region 277..298  
FT /note= "helix 12"  
FT Misc-difference 298..299  
FT /note= "helices 13-14 have been deleted from between  
FT these two sites"  
FT Region 299..342  
FT /note= "helix 15"  
XX



XX DE Human apoAIV mutein P(deltaN13, R93G).

XX KW apolipoprotein AIV; apoAIV; hypercholesterolaemia; atherosclerosis;

KW coronary disease; chylomicron; cholesterol transport; plaque deposition;

KW lecithin-cholesterol-acyltransferase; LCAT.

XX OS Synthetic.

XX Key Location/Qualifiers

FT Protein 1. .363

FT /label= (deltaN13, R93G)

FT /note= "the 13 N-terminal amino acids of human apoAIV

FT have been deleted and wild-type Arg93 has been substd. by

FT Gly"

FT Misc-difference 79

FT /note= "R93G"

XX KW09315198-A1.

XX PD 05-AUG-1993.

XX PF 26-JAN-1993; 93WO-FR0000073.

XX PR 27-JAN-1992; 92FR-00000806.

XX PA (RHON ) RHONE POULENC RORER SA.

XX PI Denefle P, Guinet F, Latta M, Murry-Brelrier A;

XX WPI; 1993-258676/32.

XX New apo:lipoprotein AIV mutant polypeptide(s) - useful for drug design

PT and/or treatment of hypercholesterolaemia.

XX Claim 7; Page 31-32 and Page 4; 42pp; French.

XX Human apolipoprotein AIV is a major component of chylomicrons in lymph

CC but is mainly found in the plasma in unassociated form. The apoAIV

CC mediates inverse transport of cholesterol. The invention covers

CC polypeptides which are derived from apoAIV by deletion of at least 10

CC terminal amino acids, by deletion of a helix or pair of helices, by

CC addition of a heterologous polypeptide portion or by a point mutation.

CC Such muteins are useful in the design of hypocholesterolaemic drugs to

CC treat hypercholesterolaemia and atherosclerosis. AAR39478 is a

CC specifically claimed mutein and the sequence has been compiled from the

CC wild-type sequence (see AAR39443) and the description given in the

CC disclosure; the mutant sequence is not shown in the specification.

CC (Updated on 25-MAR-2003 to correct PN field.)

XX SQ Sequence 363 AA;

Query Match 100.0%; Score 66; DB 2; Length 363;

Best Local Similarity 100.0%; Pred. No. 0.00085;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLVPFATLHERL 13

Db 46 KLVPFATLHERL 58

RESULT 22

ID AAR39479 standard; protein; 363 AA.

AC AAR39479;

XX 25-MAR-2003 (revised)

DT 28-JAN-1994 (first entry)

XX Human apoAIV mutein P(deltaN13).

XX apolipoprotein AIV; apoAIV; hypercholesterolaemia; atherosclerosis;

XX coronary disease; chylomicron; cholesterol transport; plaque deposition;

XX lecithin-cholesterol-acyltransferase; LCAT; tag purification marker.

XX OS Synthetic.

XX Key Location/Qualifiers

FT Peptide 1. .10

KW coronary disease; chylomicron; cholesterol transport; plaque deposition;

XX lecithin-cholesterol-acyltransferase; LCAT.

XX OS Synthetic.

XX Key Location/Qualifiers

FT Protein 1. .363

FT /label= P(deltaN13)

FT /note= "the 13 N-terminal amino acids of mature human

FT apoAIV have been deleted"

XX KW09315198-A1.

XX PD 05-AUG-1993.

XX PF 26-JAN-1993; 93WO-FR0000073.

XX PR 27-JAN-1992; 92FR-00000806.

XX PA (RHON ) RHONE POULENC RORER SA.

XX PI Denefle P, Guinet F, Latta M, Murry-Brelrier A;

XX WPI; 1993-258676/32.

XX New apo:lipoprotein AIV mutant polypeptide(s) - useful for drug design

PT and/or treatment of hypercholesterolaemia.

XX Claim 7; Page 31-32 and Page 4; 42pp; French.

XX Human apolipoprotein AIV is a major component of chylomicrons in lymph

CC but is mainly found in the plasma in unassociated form. The apoAIV

CC mediates inverse transport of cholesterol. The invention covers

CC polypeptides which are derived from apoAIV by deletion of at least 10

CC terminal amino acids, by deletion of a helix or pair of helices, by

CC addition of a heterologous polypeptide portion or by a point mutation.

CC Such muteins are useful in the design of hypocholesterolaemic drugs to

CC treat hypercholesterolaemia and atherosclerosis. AAR39479 is a

CC specifically claimed mutein and the sequence has been compiled from the

CC wild-type sequence (see AAR39443) and the description given in the

CC disclosure; the mutant sequence is not shown in the specification.

CC (Updated on 25-MAR-2003 to correct PN field.)

XX SQ Sequence 363 AA;

Query Match 100.0%; Score 66; DB 2; Length 363;

Best Local Similarity 100.0%; Pred. No. 0.00085;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLVPFATLHERL 13

Db 46 KLVPFATLHERL 58

RESULT 23

ID AAR39486 standard; protein; 373 AA.

XX AAR39486;

XX 25-MAR-2003 (revised)

DT 28-JAN-1994 (first entry)

XX Human apoAIV mutein P(tag-deltaN13).

XX apolipoprotein AIV; apoAIV; hypercholesterolaemia; atherosclerosis;

XX coronary disease; chylomicron; cholesterol transport; plaque deposition;

XX lecithin-cholesterol-acyltransferase; LCAT; tag purification marker.

XX OS Synthetic.

XX Key Location/Qualifiers

FT Peptide 1. .10



RESULT 25  
AAR39443  
ID AAR39443 standard; protein; 377 AA.  
XX AC AAR39443;  
XX XX  
DT 25-MAR-2003 (revised)  
DT 28-JAN-1994 (first entry)  
XX XX  
DE Human apolipoprotein AIV.  
XX XX  
KW apolipoprotein AIV; apoAIV; hypercholesterolaemia; atherosclerosis;  
KW coronary disease; chylomicron; cholesterol transport; plaque deposition;  
KW lecithin-cholesterol-acyltransferase; LCAT.  
XX XX  
OS Homo sapiens.  
XX XX  
FH Key Location/Qualifiers  
FT Region 14..40  
FT /note= "helix 1"  
FT Region 41..62  
FT /note= "helix 2"  
FT Region 63..95  
FT /note= "helix 3"  
FT Region 96..117  
FT /note= "helix 4"  
FT Region 118..139  
FT /note= "helix 5"  
FT Region 140..161  
FT /note= "helix 6"  
FT Region 162..183  
FT /note= "helix 7"  
FT Region 184..205  
FT /note= "helix 8"  
FT Region 206..227  
FT /note= "helix 9"  
FT Region 228..249  
FT /note= "helix 10"  
FT Region 250..267  
FT /note= "helix 11"  
FT Region 268..289  
FT /note= "helix 12"  
FT Region 290..311  
FT /note= "helix 13"  
FT Region 312..333  
FT /note= "helix 14"  
FT Region 334..377  
FT /note= "helix 15"  
XX WO9315198-A1.  
XX XX  
PD 05-AUG-1993.  
XX XX  
PF 26-JAN-1993; 93WO-FR0000073.  
XX XX  
PR 27-JAN-1992; 92FR-00000806.  
XX XX  
PA (RHON ) RHONE POULENC RORER SA.  
XX XX  
PI Deneffe P, Guinet F, Latta M, Murry-Brelier A;  
XX XX  
DR WPI; 1993-258676/32.  
DR N-PSDB; AQA43306.  
XX XX  
PT New apo:lipoprotein AIV mutant polypeptide(s) - useful for drug design  
PT and/or treatment of hypercholesterolaemia.  
XX XX  
PS Disclosure; Page 32-33; 42pp; French.  
XX XX  
CC Human apolipoprotein AIV is a major component of chylomicrons in lymph  
CC but is mainly found in the plasma in unassociated form. The apo AIV  
CC mediates inverse transport of cholesterol. The invention covers

CC polypeptides which are derived from apoAIV by deletion of at least 10  
CC terminal amino acids, by deletion of a helix or pair of helices, by  
CC addition of a heterologous polypeptide portion or by a point mutation.  
CC Such muteins are useful in the design of hypocholesterolaemic drugs to  
CC treat hypercholesterolaemia and atherosclerosis. (Updated on 25-MAR-2003  
CC to correct FN field.)  
XX XX  
SQ Sequence 377 AA;  
Query Match 100.0%; Score 66; DB 2; Length 377;  
Best Local Similarity 100.0%; Pred. No. 0.00089;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 KLVPPFATELHERL 13  
DB 60 KLVPPFATELHERL 72  
RESULT 26  
AAR45244  
ID AAR45244 standard; protein; 377 AA.  
XX AC AAR45244;  
XX XX  
DT 25-MAR-2003 (revised)  
DT 28-JAN-1994 (first entry)  
XX XX  
DE Human apoAIV mutein P(E230K).  
XX XX  
KW apolipoprotein AIV; apoAIV; hypercholesterolaemia; atherosclerosis;  
KW coronary disease; chylomicron; cholesterol transport; plaque deposition;  
KW lecithin-cholesterol-acyltransferase; LCAT.  
XX XX  
OS Synthetic.  
XX XX  
FH Key Location/Qualifiers  
FT Protein 2..377  
FT /note= "mature P(E230K)"  
FT Region 14..40  
FT /note= "helix 1"  
FT Region 41..62  
FT /note= "helix 2"  
FT Region 63..95  
FT /note= "helix 3"  
FT Region 96..117  
FT /note= "helix 4"  
FT Region 118..139  
FT /note= "helix 5"  
FT Region 140..161  
FT /note= "helix 6"  
FT Region 162..183  
FT /note= "helix 7"  
FT Region 184..205  
FT /note= "helix 8"  
FT Region 206..227  
FT /note= "helix 9"  
FT Region 228..249  
FT /note= "helix 10"  
FT Misc-difference 231  
FT /note= "wild-type Glu230 is substd. by Lys"  
FT Region 250..267  
FT /note= "helix 11"  
FT Region 268..289  
FT /note= "helix 12"  
FT Region 290..311  
FT /note= "helix 13"  
FT Region 312..333  
FT /note= "helix 14"  
FT Region 334..377  
FT /note= "helix 15"  
XX WO9315198-A1.  
XX PN  
XX XX

```

PD 05-AUG-1993.
XX
XX
XX 26-JAN-1993; 93WO-FR000073.
XX
XX 27-JAN-1992; 92FR-00000806.
XX
XX (RHON ) RHONE POULENC RORER SA.
XX
XX Denefle P, Guinet F, Latta M, Murry-Brelrier A;
XX WPI; 1993-258676/32.
XX
XX New apo.lipoprotein AIV mutant polypeptide(s) - useful for drug design
XX and/or treatment of hypercholesterolaemia.
XX
XX Claim 7; Page 31-32 and Page 5; 42pp; French.
XX
XX Human apolipoprotein AIV is a major component of chylomicrons in lymph
XX but is mainly found in the plasma in unassociated form. The apo AIV
XX mediates inverse transport of cholesterol. The invention covers
XX polypeptides which are derived from apoAIV by deletion of at least 10
XX terminal amino acids, by deletion of a helix or pair of helices, by
XX addition of a heterologous polypeptide portion or by a point mutation.
XX Such muteins are useful in the design of hypocholesterolaemic drugs to
XX treat hypercholesterolaemia and atherosclerosis. AAR45243 is a
XX specifically claimed mutein and the sequence has been compiled from the
XX wild-type sequence (see AAR39443) and the description given in the
XX disclosure; the mutant sequence is not shown in the specification.
XX (Updated on 25-MAR-2003 to correct PN field.)
XX
XX SQ Sequence 377 AA;
XX
XX Query Match 100.0%; Score 66; DB 2; Length 377;
XX Best Local Similarity 100.0%; Pred. No. 0.00089;
XX Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 KLVPPFATLHERL 13
XX |||||
XX 60 KLVPPFATLHERL 72
XX
XX RESULT 27
XX AAR39501
XX ID AAR39501 standard; protein; 377 AA.
XX
XX AC AAR39501;
XX
XX 25-MAR-2003 (revised)
XX 28-JAN-1994 (first entry)
XX
XX DE Human apoAIV mutein P(D5S).
XX
XX apolipoprotein AIV; apoAIV; hypercholesterolaemia; atherosclerosis;
XX coronary disease; chylomicron; cholesterol transport; plaque deposition;
XX lecithin-cholesterol-acyltransferase; LCAT.
XX
XX OS Synthetic.
XX
XX FH Key Location/Qualifiers
XX FT Protein 2..377
XX FT /note= "mature P(D5S)"
XX
XX Misc-difference 6
XX FT /note= "wild-type Asp5 is substd. by Ser"
XX FT 14..40
XX FT /note= "helix 1"
XX FT 41..62
XX FT /note= "helix 2"
XX FT 63..95
XX FT /note= "helix 3"
XX FT 96..117
XX FT /note= "helix 4"
XX FT 118..139
XX FT /note= "helix 5"
XX

```

```

FT Region 140..161
FT /note= "helix 6"
FT Region 162..183
FT /note= "helix 7"
FT Region 184..205
FT /note= "helix 8"
FT Region 206..227
FT /note= "helix 9"
FT Region 228..249
FT /note= "helix 10"
FT Region 250..267
FT /note= "helix 11"
FT Region 268..289
FT /note= "helix 12"
FT Region 290..311
FT /note= "helix 13"
FT Region 312..333
FT /note= "helix 14"
FT Region 334..377
FT /note= "helix 15"
XX
XX WO9315198-A1.
XX
XX 05-AUG-1993.
XX
XX 26-JAN-1993; 93WO-FR000073.
XX
XX 27-JAN-1992; 92FR-00000806.
XX
XX (RHON ) RHONE POULENC RORER SA.
XX
XX Denefle P, Guinet F, Latta M, Murry-Brelrier A;
XX WPI; 1993-258676/32.
XX
XX New apo.lipoprotein AIV mutant polypeptide(s) - useful for drug design
XX and/or treatment of hypercholesterolaemia.
XX
XX Claim 7; Page 31-32 and Page 5; 42pp; French.
XX
XX Human apolipoprotein AIV is a major component of chylomicrons in lymph
XX but is mainly found in the plasma in unassociated form. The apo AIV
XX mediates inverse transport of cholesterol. The invention covers
XX polypeptides which are derived from apoAIV by deletion of at least 10
XX terminal amino acids, by deletion of a helix or pair of helices, by
XX addition of a heterologous polypeptide portion or by a point mutation.
XX Such muteins are useful in the design of hypocholesterolaemic drugs to
XX treat hypercholesterolaemia and atherosclerosis. AAR39501 is a
XX specifically claimed mutein and the sequence has been compiled from the
XX wild-type sequence (see AAR39443) and the description given in the
XX disclosure; the mutant sequence is not shown in the specification.
XX (Updated on 25-MAR-2003 to correct PN field.)
XX
XX SQ Sequence 377 AA;
XX
XX Query Match 100.0%; Score 66; DB 2; Length 377;
XX Best Local Similarity 100.0%; Pred. No. 0.00089;
XX Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 KLVPPFATLHERL 13
XX |||||
XX 60 KLVPPFATLHERL 72
XX
XX Db
XX
XX RESULT 28
XX AAR45242
XX ID AAR45242 standard; protein; 377 AA.
XX
XX AC AAR45242;
XX
XX 25-MAR-2003 (revised)
XX 28-JAN-1994 (first entry)
XX
XX

```

DE Human apoAIV mutein P(K178Y).

XX apolipoprotein AIV; apoAIV; hypercholesterolaemia; atherosclerosis;

KW coronary disease; chylomicron; cholesterol transport; plaque deposition;

KW lecithin-cholesterol-acyltransferase; LCAT.

OS Synthetic.

XX

PH Location/Qualifiers

FT 2. .377

FT /note= "mature P(K178Y)"

FT 14. .40

FT /note= "helix 1"

FT 41. .62

FT /note= "helix 2"

FT 63. .95

FT /note= "helix 3"

FT 96. .117

FT /note= "helix 4"

FT 118. .139

FT /note= "helix 5"

FT 140. .161

FT /note= "helix 6"

FT 162. .183

FT /note= "helix 7"

FT Misc-difference 179

FT /note= "wild-type Lys178 is substd. by Tyr"

FT 184. .205

FT /note= "helix 8"

FT 206. .227

FT /note= "helix 9"

FT 228. .249

FT /note= "helix 10"

FT 250. .267

FT /note= "helix 11"

FT 268. .289

FT /note= "helix 12"

FT 290. .311

FT /note= "helix 13"

FT 312. .333

FT /note= "helix 14"

FT 334. .377

FT /note= "helix 15"

XX

FN W09315198-A1.

XX

PD 05-AUG-1993.

XX

PF 26-JAN-1993; 93WO-FR0000073.

XX

PR 27-JAN-1992; 92FR-00000806.

XX

PA (RHON ) RHONE POULENC RORER SA.

XX

PI Deneffe P, Guinet F, Latta M, Murry-Brelrier A;

XX

DR WPI; 1993-258676/32.

XX

XX New apo:lipoprotein AIV mutant polypeptide(s) - useful for drug design

FT and/or treatment of hypercholesterolaemia.

XX

PS Claim 7; Page 31-32 and Page 5; 42pp; French.

XX

CC Human apolipoprotein AIV is a major component of chylomicrons in lymph

CC but is mainly found in the plasma in unassociated form. The apo AIV

CC mediates inverse transport of cholesterol. The invention covers

CC polypeptides which are derived from apoAIV by deletion of at least 10

CC terminal amino acids, by deletion of a helix or pair of helices, by

CC addition of a heterologous polypeptide portion or by a point mutation.

CC Such mutants are useful in the design of hypocholesterolaemic drugs to

CC treat hypercholesterolaemia and atherosclerosis. AAR39480 is a

CC specifically claimed mutein and the sequence has been compiled from the

CC wild-type sequence (see AAR39443) and the description given in the

CC disclosure; the mutant sequence is not shown in the specification.

CC (Updated on 25-MAR-2003 to correct PN field.)

XX

SQ Sequence 377 AA;

CC disclosure; the mutant sequence is not shown in the specification.

CC (Updated on 25-MAR-2003 to correct PN field.)

XX

SQ Sequence 377 AA;

Query Match 100.0%; Score 66; DB 2; Length 377;

Best Local Similarity 100.0%; Pred. No. 0.00089;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KLVPEATELHERL 13

DB 60 KLVPEATELHERL 72

RESULT 29

AAR39480

ID AAR39480 standard; protein; 377 AA.

XX

AC AAR39480;

XX

XX 25-MAR-2003 (revised)

DT 28-JAN-1994 (first entry)

XX

XX Human apoAIV mutein P(R93G).

DE

XX

KW apolipoprotein AIV; apoAIV; hypercholesterolaemia; atherosclerosis;

KW coronary disease; chylomicron; cholesterol transport; plaque deposition;

KW lecithin-cholesterol-acyltransferase; LCAT.

XX

OS Synthetic.

XX

PH Key Location/Qualifiers

FT Protein 2. .377

FT /note= "mature P(R93G)"

FT Misc-difference 94

FT /note= "wild-type Arg93 is substd. by Gly"

FT

XX

FN W09315198-A1.

XX

PD 05-AUG-1993.

XX

PF 26-JAN-1993; 93WO-FR0000073.

XX

PR 27-JAN-1992; 92FR-00000806.

XX

PA (RHON ) RHONE POULENC RORER SA.

XX

PI Deneffe P, Guinet F, Latta M, Murry-Brelrier A;

XX

DR WPI; 1993-258676/32.

XX

XX New apo:lipoprotein AIV mutant polypeptide(s) - useful for drug design

FT and/or treatment of hypercholesterolaemia.

XX

PS Claim 7; Page 31-32 and Page 4; 42pp; French.

XX

CC Human apolipoprotein AIV is a major component of chylomicrons in lymph

CC but is mainly found in the plasma in unassociated form. The apoAIV

CC mediates inverse transport of cholesterol. The invention covers

CC polypeptides which are derived from apoAIV by deletion of at least 10

CC terminal amino acids, by deletion of a helix or pair of helices, by

CC addition of a heterologous polypeptide portion or by a point mutation.

CC Such mutants are useful in the design of hypocholesterolaemic drugs to

CC treat hypercholesterolaemia and atherosclerosis. AAR39480 is a

CC specifically claimed mutein and the sequence has been compiled from the

CC wild-type sequence (see AAR39443) and the description given in the

CC disclosure; the mutant sequence is not shown in the specification.

CC (Updated on 25-MAR-2003 to correct PN field.)

XX

SQ Sequence 377 AA;

Query Match 100.0%; Score 66; DB 2; Length 377;

Best Local Similarity 100.0%; Pred. No. 0.00089;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLVPFATELHERL 13  
Db 60 KLVPFATELHERL 72

RESULT 30  
AAR39499  
ID AAR39499 standard; protein; 377 AA.  
XX AAR39499;  
XX  
XX 25-MAR-2003 (revised)  
DT 28-JAN-1994 (first entry)  
XX  
DE Human apoAIV mutein P(D44F).

XX  
KW apolipoprotein AIV; apoAIV; hypercholesterolaemia; atherosclerosis;  
KW coronary disease; chylomicron; cholesterol transport; plaque deposition;  
KW lecithin-cholesterol-acyltransferase; LCAT.  
XX  
OS Synthetic.

XX Key Location/Qualifiers  
FH Protein 2..377  
FT /note= "mature P(D44F)"  
FT Region 14..40  
FT /note= "helix 1"  
FT Region 41..62  
FT /note= "helix 2"  
FT Misc-difference 45  
FT /note= "wild-type Asp44 is substnd. by Phe"  
FT Region 63..95  
FT /note= "helix 3"  
FT Region 96..117  
FT /note= "helix 4"  
FT Region 118..139  
FT /note= "helix 5"  
FT Region 140..161  
FT /note= "helix 6"  
FT Region 162..183  
FT /note= "helix 7"  
FT Region 184..205  
FT /note= "helix 8"  
FT Region 206..227  
FT /note= "helix 9"  
FT Region 228..249  
FT /note= "helix 10"  
FT Region 250..267  
FT /note= "helix 11"  
FT Region 268..289  
FT /note= "helix 12"  
FT Region 290..311  
FT /note= "helix 13"  
FT Region 312..333  
FT /note= "helix 14"  
FT Region 334..377  
FT /note= "helix 15"

XX WO9315198-A1.  
XX  
XX  
PD 05-AUG-1993.  
XX  
XX 26-JAN-1993; 93WO-FR000073.  
XX  
XX 27-JAN-1992; 92FR-0000806.  
XX (RHON ) RHONE POULENC RORER SA.  
XX  
XX Deneffe P, Guinet F, Latta M, Murry-Brelrier A;  
XX WPI; 1993-258676/32.

XX New apo:lipoprotein AIV mutant polypeptide(s) - useful for drug design  
PT and/or treatment of hypercholesterolaemia.  
XX  
PS Claim 7; Page 31-32 and Page 5; 42pp; French.  
XX  
CC Human apolipoprotein AIV is a major component of chylomicrons in lymph  
CC but is mainly found in the plasma in unassociated form. The apo AIV  
CC mediates inverse transport of cholesterol. The invention covers  
CC polypeptides which are derived from apoAIV by deletion of at least 10  
CC terminal amino acids, by deletion of a helix or pair of helices, by  
CC addition of a heterologous polypeptide portion or by a point mutation.  
CC Such mutants are useful in the design of hypocholesterolaemic drugs to  
CC treat hypercholesterolaemia and atherosclerosis. AAR39499 is a  
CC specifically claimed mutein and the sequence has been compiled from the  
CC wild-type sequence (see AAR39443) and the description given in the  
CC disclosure; the mutant sequence is not shown in the specification.  
CC (Updated on 25-MAR-2003 to correct PN field.)  
XX  
SQ Sequence 377 AA;  
Query Match 100.0%; Score 66; DB 2; Length 377;  
Best Local Similarity 100.0%; Pred No. 0.00089;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KLVPFATELHERL 13  
Db 60 KLVPFATELHERL 72

RESULT 31  
AAR39500  
ID AAR39500 standard; protein; 377 AA.  
XX AAR39500;  
XX  
XX 25-MAR-2003 (revised)  
DT 28-JAN-1994 (first entry)  
XX  
XX Human apoAIV mutein P(D44A).  
XX  
XX apolipoprotein AIV; apoAIV; hypercholesterolaemia; atherosclerosis;  
KW coronary disease; chylomicron; cholesterol transport; plaque deposition;  
KW lecithin-cholesterol-acyltransferase; LCAT.  
XX  
OS Synthetic.

XX Key Location/Qualifiers  
FH Protein 2..377  
FT /note= "mature P(D44A)"  
FT Region 14..40  
FT /note= "helix 1"  
FT Region 41..62  
FT /note= "helix 2"  
FT Misc-difference 45  
FT /note= "wild-type Asp44 is substnd. by Ala"  
FT Region 63..95  
FT /note= "helix 3"  
FT Region 96..117  
FT /note= "helix 4"  
FT Region 118..139  
FT /note= "helix 5"  
FT Region 140..161  
FT /note= "helix 6"  
FT Region 162..183  
FT /note= "helix 7"  
FT Region 184..205  
FT /note= "helix 8"  
FT Region 206..227  
FT /note= "helix 9"  
FT Region 228..249  
FT /note= "helix 10"  
FT Region 250..267



FT Region /note= "helix 11"  
 FT 268..289  
 FT /note= "helix 12"  
 FT 290..311  
 FT /note= "helix 13"  
 FT 312..333  
 FT /note= "helix 14"  
 FT 334..377  
 FT /note= "helix 15"

XX WO9315198-A1.

XX 05-AUG-1993.

XX 26-JAN-1993; 93WO-FR0000073.

XX 27-JAN-1992; 92FR-00000806.

XX (RHON ) RHONE POULENC RORER SA.

XX Denefle P, Guinet F, Latta M, Murry-Brelrier A;

XX WPI; 1993-258676/32.

XX New apo:lipoprotein AIV mutant polypeptide(s) - useful for drug design  
 and/or treatment of hypercholesterolaemia.

XX Claim 7; Page 31-32 and Page 5; 42pp; French.

XX Human apolipoprotein AIV is a major component of chylomicrons in lymph  
 but is mainly found in the plasma in unassociated form. The apo AIV  
 mediates inverse transport of cholesterol. The invention covers  
 polypeptides which are derived from apoAIV by deletion of at least 10  
 terminal amino acids, by deletion of a helix or pair of helices, by  
 addition of a heterologous polypeptide portion or by a point mutation.  
 Such muteins are useful in the design of hypocholesterolaemic drugs to  
 treat hypercholesterolaemia and atherosclerosis. AAR39500 is a  
 specifically claimed mutein and the sequence has been compiled from the  
 wild-type sequence (see AAR39443) and the description given in the  
 disclosure; the mutant sequence is not shown in the specification.  
 (Updated on 25-MAR-2003 to correct PN field.)

XX Sequence 377 AA;

Query Match 100.0%; Score 66; DB 2; Length 377;

Best Local Similarity 100.0%; Pred. No. 0.00089;  
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLVPFATELHERL 13

Db 60 KLVPFATELHERL 72

RESULT 32

AAR45243

ID AAR45243 standard; protein; 377 AA.

XX AAR45243;

XX 25-MAR-2003 (revised)

DT 28-JAN-1994 (first entry)

XX Human apoAIV mutein P(K178A).

XX apolipoprotein AIV; apoAIV; hypercholesterolaemia; atherosclerosis;  
 KW coronary disease; chylomicron; cholesterol transport; plaque deposition;  
 KW lecithin-cholesterol-acyltransferase; LCAT.

XX Synthetic.

XX Key Location/Qualifiers

FT Protein 2..377

FT /note= "mature P(K178A)"

FT Region 14..40  
 FT /note= "helix 1"  
 FT 41..62  
 FT /note= "helix 2"  
 FT 63..95  
 FT /note= "helix 3"  
 FT 96..117  
 FT /note= "helix 4"  
 FT 118..139  
 FT /note= "helix 5"  
 FT 140..161  
 FT /note= "helix 6"  
 FT 162..183  
 FT /note= "helix 7"  
 FT Misc-difference 179  
 FT /note= "wild-type Lys178 is substd. by Ala"  
 FT 184..205  
 FT /note= "helix 8"  
 FT 206..227  
 FT /note= "helix 9"  
 FT 228..249  
 FT /note= "helix 10"  
 FT 250..267  
 FT /note= "helix 11"  
 FT 268..289  
 FT /note= "helix 12"  
 FT 290..311  
 FT /note= "helix 13"  
 FT 312..333  
 FT /note= "helix 14"  
 FT 334..377  
 FT /note= "helix 15"

XX WO9315198-A1.

XX 05-AUG-1993.

XX 26-JAN-1993; 93WO-FR0000073.

XX 27-JAN-1992; 92FR-00000806.

XX (RHON ) RHONE POULENC RORER SA.

XX Denefle P, Guinet F, Latta M, Murry-Brelrier A;

XX WPI; 1993-258676/32.

XX New apo:lipoprotein AIV mutant polypeptide(s) - useful for drug design  
 and/or treatment of hypercholesterolaemia.

XX Claim 7; Page 31-32 and Page 5; 42pp; French.

XX Human apolipoprotein AIV is a major component of chylomicrons in lymph  
 but is mainly found in the plasma in unassociated form. The apo AIV  
 mediates inverse transport of cholesterol. The invention covers  
 polypeptides which are derived from apoAIV by deletion of at least 10  
 terminal amino acids, by deletion of a helix or pair of helices, by  
 addition of a heterologous polypeptide portion or by a point mutation.  
 Such muteins are useful in the design of hypocholesterolaemic drugs to  
 treat hypercholesterolaemia and atherosclerosis. AAR45243 is a  
 specifically claimed mutein and the sequence has been compiled from the  
 wild-type sequence (see AAR39443) and the description given in the  
 disclosure; the mutant sequence is not shown in the specification.  
 (Updated on 25-MAR-2003 to correct PN field.)

XX Sequence 377 AA;

Query Match 100.0%; Score 66; DB 2; Length 377;

Best Local Similarity 100.0%; Pred. No. 0.00089;  
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLVPFATELHERL 13

Db 60 KLVPFATELHERL 72

Db 60 KLVFFATELHERL 72

RESULT 33  
 AAB90663  
 ID AAB90663 standard; protein; 382 AA.  
 XX  
 AC AAB90663;  
 XX  
 DT 01-JUN-2001 (first entry)  
 XX  
 DE Human secreted protein, SEQ ID NO: 206.  
 XX  
 KW Human; secreted protein; immunomodulatory; antisclerotic; dermatological;  
 KW anti-inflammatory; anti-HIV; cytostatic; cardiant; vascular;  
 KW anti-angiogenic; ophthalmological; neuroprotectant; neurotropic;  
 KW anticonvulsant; antialzheimers; antiparkinsonian; antimicrobial;  
 KW vulnary; vaccine; gene therapy; cancer; protein coordinate data;  
 KW infection.  
 XX  
 OS Homo sapiens.  
 XX WO200121658-A1.  
 XX  
 XX 29-MAR-2001.  
 XX  
 XX 22-SEP-2000; 2000WO-US026013.  
 XX  
 XX 24-SEP-1999; 99US-0155709P.  
 XX  
 XX (HUMA-) HUMAN GENOME SCI INC.  
 XX  
 XX Ni J, Baker KP, Birse CE, Ebner R, Fiscella M, Komatsoulis GA;  
 XX Lafleur DW, Moore PA, Olsen HS, Rosen CA, Ruben SA, Soppet DR;  
 XX Young PE, Wei P, Florence KA;  
 XX WPI; 2001-235311/24.  
 XX  
 XX Nucleic acids encoding 32 human secreted polypeptides, useful for  
 XX preventing, diagnosing and/or treating e.g. cancers, Parkinson's disease  
 XX and diabetic retinopathy.  
 XX  
 XX Disclosure; Page 869-870; 890pp; English.  
 XX  
 XX The present sequence is provided in a specification relating to nucleic  
 XX acid molecules encoding 32 novel human secreted polypeptides. The nucleic  
 XX acid molecules and polypeptides may be used in the prevention, diagnosis  
 XX and treatment of diseases such as immune disorders (e.g. multiple  
 XX sclerosis, systemic lupus erythematosus and human immuno-deficiency virus  
 XX (HIV) infections), hyperproliferative disorders (e.g. cancers and  
 XX Gaucher's disease), cardiovascular diseases (e.g. Scimitar syndrome,  
 XX Chaga's cardiomyopathy and coronary arteriosclerosis), angiogenic  
 XX disorders (e.g. corneal graft neovascularisation and diabetic  
 XX retinopathy), neurological disorders (e.g. Huntington's chorea,  
 XX Alzheimer's disease and Parkinson's disease), infectious diseases and/or  
 XX for promoting wound healing, regeneration and/or chemotaxis. The nucleic  
 XX acid molecules may be used to produce the secreted polypeptides. They may  
 XX also be used as DNA probes in diagnostic assays to detect and quantitate  
 XX the presence of similar nucleic acid sequences in samples. The  
 XX polypeptides may be used as antigens in the production of antibodies and  
 XX in assays to identify modulators of their expression and activity  
 XX  
 XX Sequence 382 AA;  
 XX  
 XX Query Match 100.0%; Score 66; DB 4; Length 382;  
 XX Best Local Similarity 100.0%; Pred. No. 0.00091;  
 XX Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 XX  
 QY 1 KLVFFATELHERL 13  
 DB 79 KLVFFATELHERL 91

RESULT 34  
 AAB90664  
 ID AAB90664 standard; protein; 396 AA.  
 XX  
 AC AAB90664;  
 XX  
 DT 01-JUN-2001 (first entry)  
 XX  
 DE Human secreted protein, SEQ ID NO: 207.  
 XX  
 KW Human; secreted protein; immunomodulatory; antisclerotic; dermatological;  
 KW anti-inflammatory; anti-HIV; cytostatic; cardiant; vascular;  
 KW anti-angiogenic; ophthalmological; neuroprotectant; neurotropic;  
 KW anticonvulsant; antialzheimers; antiparkinsonian; antimicrobial;  
 KW vulnary; vaccine; gene therapy; cancer; protein coordinate data;  
 KW infection.  
 XX  
 OS Homo sapiens.  
 XX WO200121658-A1.  
 XX  
 XX 29-MAR-2001.  
 XX  
 XX 22-SEP-2000; 2000WO-US026013.  
 XX  
 XX 24-SEP-1999; 99US-0155709P.  
 XX  
 XX (HUMA-) HUMAN GENOME SCI INC.  
 XX  
 XX Ni J, Baker KP, Birse CE, Ebner R, Fiscella M, Komatsoulis GA;  
 XX Lafleur DW, Moore PA, Olsen HS, Rosen CA, Ruben SA, Soppet DR;  
 XX Young PE, Wei P, Florence KA;  
 XX WPI; 2001-235311/24.  
 XX  
 XX Nucleic acids encoding 32 human secreted polypeptides, useful for  
 XX preventing, diagnosing and/or treating e.g. cancers, Parkinson's disease  
 XX and diabetic retinopathy.  
 XX  
 XX Disclosure; Page 870-871; 890pp; English.  
 XX  
 XX The present sequence is provided in a specification relating to nucleic  
 XX acid molecules encoding 32 novel human secreted polypeptides. The nucleic  
 XX acid molecules and polypeptides may be used in the prevention, diagnosis  
 XX and treatment of diseases such as immune disorders (e.g. multiple  
 XX sclerosis, systemic lupus erythematosus and human immuno-deficiency virus  
 XX (HIV) infections), hyperproliferative disorders (e.g. cancers and  
 XX Gaucher's disease), cardiovascular diseases (e.g. Scimitar syndrome,  
 XX Chaga's cardiomyopathy and coronary arteriosclerosis), angiogenic  
 XX disorders (e.g. corneal graft neovascularisation and diabetic  
 XX retinopathy), neurological disorders (e.g. Huntington's chorea,  
 XX Alzheimer's disease and Parkinson's disease), infectious diseases and/or  
 XX for promoting wound healing, regeneration and/or chemotaxis. The nucleic  
 XX acid molecules may be used to produce the secreted polypeptides. They may  
 XX also be used as DNA probes in diagnostic assays to detect and quantitate  
 XX the presence of similar nucleic acid sequences in samples. The  
 XX polypeptides may be used as antigens in the production of antibodies and  
 XX in assays to identify modulators of their expression and activity  
 XX  
 XX Sequence 396 AA;  
 XX  
 XX Query Match 100.0%; Score 66; DB 4; Length 396;  
 XX Best Local Similarity 100.0%; Pred. No. 0.00094;  
 XX Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 XX  
 QY 1 KLVFFATELHERL 13  
 DB 79 KLVFFATELHERL 91

RESULT 35  
 AAO15885  
 ID AAO15885 standard; protein; 396 AA.

XX AAO15885;  
 AC  
 XX  
 DT 30-JAN-2003 (first entry)  
 XX  
 DE Human apolipoprotein A-IV (APOA4).  
 XX  
 XX Human; gene therapy; single nucleotide polymorphism; SNP;  
 KW cytochrome C oxidase subunit VIB; COX6B; high serum cholesterol; GPI-1;  
 KW N-acetylglucosaminyl transferase component; cardiovascular disease; HDL;  
 KW glycosylphosphatidylinositol-1; low serum high density lipoprotein.  
 XX  
 XX Homo sapiens.  
 OS  
 XX  
 XX WQ200272604-A2.  
 PN  
 XX  
 PD 19-SEP-2002.  
 XX  
 XX 05-MAR-2002; 2002WO-US006728.  
 XX  
 XX 09-MAR-2001; 2001US-00802640.  
 PR  
 XX (SEQU-) SEQUENOM INC.  
 PA  
 XX Braun A, Bansal A, Kleyl PW;  
 PI  
 XX WPI; 2002-750478/81.  
 DR  
 DR N-PSDB; ABT13007.  
 XX  
 XX Detecting the presence or absence of an allelic variant of a polymorphic  
 PT region of COX6B and/or GPI-1 gene, useful for detecting a predisposition  
 PT to high serum cholesterol, low serum HDL and cardiovascular disease.  
 XX  
 XX Disclosure; Page 117; 199pp; English.  
 PS  
 XX The invention comprises methods of detecting the presence or absence of  
 CC at least one allelic variant of a polymorphic region of a gene associated  
 CC with cardiovascular disease. The invention specifically relates to  
 CC detecting the region of a cytochrome C oxidase subunit VIB (COX6B) gene  
 CC that is associated with high serum cholesterol, or the region of the N-  
 CC acetylglucosaminyl transferase component glycosylphosphatidylinositol-1  
 CC (GPI-1) gene that is associated with low serum high density lipoprotein  
 CC (HDL). The methods of the invention are useful for detecting a  
 CC predisposition to high serum cholesterol, low serum HDL and  
 CC cardiovascular disease. The methods are also useful for elucidating  
 CC pathological pathways, developing diagnostic assays and new drug  
 CC therapies for such disorders. The present amino acid sequence represents  
 CC a human protein that is associated with high serum cholesterol, low serum  
 CC HDL and/or cardiovascular disease  
 XX  
 SQ Sequence 396 AA;  
 Query Match 100.0%; Score 66; DB 5; Length 396;  
 Best Local Similarity 100.0%; Pred. No. 0.00094;  
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 KLVPFATLHERL 13  
 DB |||||  
 79 KLVPFATLHERL 91  
 RESULT 36  
 AAU10862  
 ID AAU10862 standard; protein; 396 AA.  
 XX  
 AC AAU10862;  
 XX  
 DT 14-FEB-2002 (first entry)  
 XX  
 XX Human apolipoprotein A-IV, APOA4, isoform #2.  
 DE  
 XX Human; APOA4; apolipoprotein A-IV; antiatherosclerotic; cardiant;  
 KW haplotype; chromosome 11q23-qter; coronary heart disease; obesity;  
 KW

atherosclerosis.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT Misc-difference 111 /note= "Wild-type Ala substituted by Val"  
 XX  
 FT WQ200177124-A2.  
 PN  
 XX 18-OCT-2001.  
 PD  
 XX 03-APR-2001; 2001WO-US010670.  
 PF  
 XX 05-APR-2000; 2000US-0194362P.  
 PR  
 XX (GENA-) GENAISSANCE PHARM INC.  
 PA  
 XX Bentivegna SC, Choi JY, Kilem SE, Koshy B;  
 PI WPI; 2002-041281/05.  
 XX  
 DR New haplotypes of the human apolipoprotein A-IV gene, useful to diagnose  
 PT and treat disorders associated with its abnormal expression or function  
 PT such as coronary artery disease.  
 XX  
 PS Claim 29; Page; 71pp; English.  
 XX  
 CC The invention relates to haplotyping the human apolipoprotein A-IV  
 CC (APOA4) gene of an individual, comprising determining if the individual  
 CC has one of the APOA4 haplotypes or haplotype pairs fully defined in the  
 CC specification. Also disclosed are genotyping oligonucleotides (or allele  
 CC specific oligonucleotides, ASO) as well as methods for correlating a  
 CC particular haplotype pair with a trait e.g. obesity, in a population. The  
 CC APOA4 gene is located on chromosome 11q23-qter. The methods of the  
 CC invention are useful to diagnose and develop treatment for disorders  
 CC associated with abnormal APOA4 expression or function, for example  
 CC coronary heart disease and atherosclerosis. The APOA4 isogenes and  
 CC screened compounds are useful for the treatment of disorders associated  
 CC with abnormal APOA4 expression or function such as coronary artery  
 CC disease. The present sequence is an isoform of the APOA4 protein. Note:  
 CC The present sequence is not shown in the specification but was created by  
 CC the indexer from the APOA4 sequence shown in figure 3  
 XX  
 SQ Sequence 396 AA;  
 Query Match 100.0%; Score 66; DB 5; Length 396;  
 Best Local Similarity 100.0%; Pred. No. 0.00094;  
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 KLVPFATLHERL 13  
 DB |||||  
 79 KLVPFATLHERL 91  
 RESULT 37  
 AAU10863  
 ID AAU10863 standard; protein; 396 AA.  
 XX  
 AC AAU10863;  
 XX  
 DT 14-FEB-2002 (first entry)  
 XX  
 XX Human apolipoprotein A-IV, APOA4, isoform #3.  
 DE  
 XX Human; APOA4; apolipoprotein A-IV; antiatherosclerotic; cardiant;  
 KW haplotype; chromosome 11q23-qter; coronary heart disease; obesity;  
 KW atherosclerosis.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT Misc-difference 127

```

FT XX /note= "Wild-type Asn substituted by Lys"
XX PN WO200177124-A2.
XX PD 18-OCT-2001.
XX XX
XX XX
XX PF 03-APR-2001; 2001WO-US010670.
XX PR 05-APR-2000; 2000US-0194362P.
XX XX (GENA-) GENAISSANCE PHARM INC.
XX PA Bentivegna SC, Choi JY, Kliem SE, Koshy B;
XX PI WPI; 2002-041281/05.
XX DR
XX XX
XX XX New haplotypes of the human apolipoprotein A-IV gene, useful to diagnose
PT and treat disorders associated with its abnormal expression or function
PT such as coronary artery disease.
XX PS
XX PS Claim 29; Page; 71pp; English.
XX CC The invention relates to haplotyping the human apolipoprotein A-IV
CC (APOA4) gene of an individual, comprising determining if the individual
CC has one of the APOA4 haplotypes or haplotype pairs fully defined in the
CC specification. Also disclosed are genotyping oligonucleotides (or allele
CC specific oligonucleotides, ASO) as well as methods for correlating a
CC particular haplotype pair with a trait e.g. obesity, in a population. The
CC APOA4 gene is located on chromosome 11q23-qter. The methods of the
CC invention are useful to diagnose and develop treatment for disorders
CC associated with abnormal APOA4 expression or function, for example
CC coronary heart disease and atherosclerosis. The APOA4 isogenes and
CC screened compounds are useful for the treatment of disorders associated
CC with abnormal APOA4 expression or function such as coronary artery
CC disease. The present sequence is an isoform of the APOA4 protein. Note:
CC The present sequence is not shown in the specification but was created by
CC the indexer from the APOA4 sequence shown in figure 3
XX SQ
SQ Sequence 396 AA;
Query Match 100.0%; Score 66; DB 5; Length 396;
Best Local Similarity 100.0%; Pred. No. 0.00094;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KLVPPFATLHERL 13
Db |||||
79 KLVPPFATLHERL 91

RESULT 38
AAU10865
ID AAU10865 standard; protein; 396 AA.
XX AC AAU10865;
XX XX
XX DT 14-FEB-2002 (first entry)
XX DE Human apolipoprotein A-IV, APOA4, isoform #5.
XX XX
XX KW Human; APOA4; apolipoprotein A-IV; antiatherosclerotic; cardiant;
XX KW haplotype; chromosome 11q23-qter; coronary heart disease; obesity;
XX KW atherosclerosis.
XX OS Homo sapiens.
XX XX
XX FH Key Location/Qualifiers
XX FT Misc-difference 161
XX FT /note= "Wild-type Ala substituted by Ser"
XX XX
XX PN WO200177124-A2.
XX PD 18-OCT-2001.
XX XX
XX XX
XX PF 03-APR-2001; 2001WO-US010670.
XX PR 05-APR-2000; 2000US-0194362P.
XX XX (GENA-) GENAISSANCE PHARM INC.
XX PA Bentivegna SC, Choi JY, Kliem SE, Koshy B;
XX PI WPI; 2002-041281/05.
XX DR
XX XX
XX XX New haplotypes of the human apolipoprotein A-IV gene, useful to diagnose
PT and treat disorders associated with its abnormal expression or function
PT such as coronary artery disease.
XX PS
XX PS Claim 29; Page; 71pp; English.
XX CC The invention relates to haplotyping the human apolipoprotein A-IV
CC (APOA4) gene of an individual, comprising determining if the individual
CC has one of the APOA4 haplotypes or haplotype pairs fully defined in the
CC specification. Also disclosed are genotyping oligonucleotides (or allele
CC specific oligonucleotides, ASO) as well as methods for correlating a
CC particular haplotype pair with a trait e.g. obesity, in a population. The
CC APOA4 gene is located on chromosome 11q23-qter. The methods of the
CC invention are useful to diagnose and develop treatment for disorders
CC associated with abnormal APOA4 expression or function, for example
CC coronary heart disease and atherosclerosis. The APOA4 isogenes and
CC screened compounds are useful for the treatment of disorders associated
CC with abnormal APOA4 expression or function such as coronary artery
CC disease. The present sequence is an isoform of the APOA4 protein. Note:
CC The present sequence is not shown in the specification but was created by
CC the indexer from the APOA4 sequence shown in figure 3
XX SQ
SQ Sequence 396 AA;
Query Match 100.0%; Score 66; DB 5; Length 396;
Best Local Similarity 100.0%; Pred. No. 0.00094;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KLVPPFATLHERL 13
Db |||||
79 KLVPPFATLHERL 91

RESULT 39
AAU10867
ID AAU10867 standard; protein; 396 AA.
XX AC AAU10867;
XX XX
XX DT 14-FEB-2002 (first entry)
XX DE Human apolipoprotein A-IV, APOA4, isoform #7.
XX XX
XX KW Human; APOA4; apolipoprotein A-IV; antiatherosclerotic; cardiant;
XX KW haplotype; chromosome 11q23-qter; coronary heart disease; obesity;
XX KW atherosclerosis.
XX OS Homo sapiens.
XX XX
XX FH Key Location/Qualifiers
XX FT Misc-difference 353
XX FT /note= "Wild-type Ser substituted by Ala"
XX XX
XX PN WO200177124-A2.
XX PD 18-OCT-2001.
XX XX
XX PF 03-APR-2001; 2001WO-US010670.
XX PR 05-APR-2000; 2000US-0194362P.
XX XX (GENA-) GENAISSANCE PHARM INC.
XX PA
XX XX

```

PI Bentivegna SC, Choi JY, Kliem SE, Koshy B;  
 XX WPI; 2002-041281/05.  
 XX  
 PT New haplotypes of the human apolipoprotein A-IV gene, useful to diagnose  
 PT and treat disorders associated with its abnormal expression or function  
 PT such as coronary artery disease.  
 XX  
 XX  
 PS Claim 29; Page; 71pp; English.  
 XX  
 CC The invention relates to haplotyping the human apolipoprotein A-IV  
 CC (APOA4) gene of an individual, comprising determining if the individual  
 CC has one of the APOA4 haplotypes or haplotype pairs fully defined in the  
 CC specification. Also disclosed are genotyping oligonucleotides (or allele  
 CC specific oligonucleotides, ASO) as well as methods for correlating a  
 CC particular haplotype pair with a trait e.g. obesity, in a population. The  
 CC APOA4 gene is located on chromosome 11q23-qter. The methods of the  
 CC invention are useful to diagnose and develop treatment for disorders  
 CC associated with abnormal APOA4 expression or function, for example  
 CC coronary heart disease and atherosclerosis. The APOA4 isogenes and  
 CC screened compounds are useful for the treatment of disorders associated  
 CC with abnormal APOA4 expression or function such as coronary artery  
 CC disease. The present sequence is an isoform of the APOA4 protein. Note:  
 CC The present sequence is not shown in the specification but was created by  
 CC the indexer from the APOA4 sequence shown in figure 3  
 XX  
 SQ Sequence 396 AA;  
 Query Match 100.0%; Score 66; DB 5; Length 396;  
 Best Local Similarity 100.0%; Pred. No. 0.00094;  
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 KLVPPFATELHERL 13  
 DB 79 KLVPPFATELHERL 91  
 RESULT 40  
 AAU10866  
 ID AAU10866 standard; protein; 396 AA.  
 AC  
 XX AAU10866;  
 XX 14-FEB-2002 (first entry)  
 XX Human apolipoprotein A-IV, APOA4, isoform #6.  
 DE  
 DE Human; APOA4; apolipoprotein A-IV; antiatherosclerotic; cardiant;  
 KW haplotype; chromosome 11q23-qter; coronary heart disease; obesity;  
 KW atherosclerosis.  
 XX Homo sapiens.  
 OS  
 XX Key Location/Qualifiers  
 FH Misc-difference 193  
 FT /label= Met, Ile  
 FT /note= "replaces wild-type val"  
 FT  
 XX WO200177124-A2.  
 XX 18-OCT-2001.  
 XX 03-APR-2001; 2001WO-US010670.  
 XX 05-APR-2000; 2000US-0194362P.  
 XX (GENA-) GENAISSANCE PHARM INC.  
 XX Bentivegna SC, Choi JY, Kliem SE, Koshy B;  
 XX WPI; 2002-041281/05.  
 XX New haplotypes of the human apolipoprotein A-IV gene, useful to diagnose

PT and treat disorders associated with its abnormal expression or function  
 PT such as coronary artery disease.  
 XX  
 PS Claim 29; Page; 71pp; English.  
 XX  
 CC The invention relates to haplotyping the human apolipoprotein A-IV  
 CC (APOA4) gene of an individual, comprising determining if the individual  
 CC has one of the APOA4 haplotypes or haplotype pairs fully defined in the  
 CC specification. Also disclosed are genotyping oligonucleotides (or allele  
 CC specific oligonucleotides, ASO) as well as methods for correlating a  
 CC particular haplotype pair with a trait e.g. obesity, in a population. The  
 CC APOA4 gene is located on chromosome 11q23-qter. The methods of the  
 CC invention are useful to diagnose and develop treatment for disorders  
 CC associated with abnormal APOA4 expression or function, for example  
 CC coronary heart disease and atherosclerosis. The APOA4 isogenes and  
 CC screened compounds are useful for the treatment of disorders associated  
 CC with abnormal APOA4 expression or function such as coronary artery  
 CC disease. The present sequence is an isoform of the APOA4 protein. Note:  
 CC The present sequence is not shown in the specification but was created by  
 CC the indexer from the APOA4 sequence shown in figure 3  
 XX  
 SQ Sequence 396 AA;  
 Query Match 100.0%; Score 66; DB 5; Length 396;  
 Best Local Similarity 100.0%; Pred. No. 0.00094;  
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 KLVPPFATELHERL 13  
 DB 79 KLVPPFATELHERL 91  
 RESULT 41  
 AAU10869  
 ID AAU10869 standard; protein; 396 AA.  
 AC  
 XX AAU10869;  
 XX 14-FEB-2002 (first entry)  
 XX Human apolipoprotein A-IV, APOA4, isoform #9.  
 DE  
 DE Human; APOA4; apolipoprotein A-IV; antiatherosclerotic; cardiant;  
 KW haplotype; chromosome 11q23-qter; coronary heart disease; obesity;  
 KW atherosclerosis.  
 XX Homo sapiens.  
 OS  
 XX Key Location/Qualifiers  
 FH Misc-difference 380  
 FT /note= "Wild-type His substituted by Gln"  
 FT  
 XX WO200177124-A2.  
 XX 18-OCT-2001.  
 XX 03-APR-2001; 2001WO-US010670.  
 XX 05-APR-2000; 2000US-0194362P.  
 XX (GENA-) GENAISSANCE PHARM INC.  
 XX Bentivegna SC, Choi JY, Kliem SE, Koshy B;  
 XX WPI; 2002-041281/05.  
 XX New haplotypes of the human apolipoprotein A-IV gene, useful to diagnose  
 PT and treat disorders associated with its abnormal expression or function  
 PT such as coronary artery disease.  
 XX  
 XX Claim 29; Page; 71pp; English.  
 XX The invention relates to haplotyping the human apolipoprotein A-IV

CC (APOA4) gene of an individual, comprising determining if the individual  
 CC has one of the APOA4 haplotypes or haplotype pairs fully defined in the  
 CC specification. Also disclosed are genotyping oligonucleotides (or allele  
 CC specific oligonucleotides, ASO) as well as methods for correlating a  
 CC particular haplotype pair with a trait e.g. obesity, in a population. The  
 CC APOA4 gene is located on chromosome 11q23-qter. The methods of the  
 CC invention are useful to diagnose and develop treatment for disorders  
 CC associated with abnormal APOA4 expression or function, for example  
 CC coronary heart disease and atherosclerosis. The APOA4 isogenes and  
 CC screened compounds are useful for the treatment of disorders associated  
 CC with abnormal APOA4 expression or function such as coronary artery  
 CC disease. The present sequence is an isoform of the APOA4 protein. Note:  
 CC The present sequence is not shown in the specification but was created by  
 CC the indexer from the APOA4 sequence shown in figure 3

XX SQ Sequence 396 AA;

Query Match 100.0%; Score 66; DB 5; Length 396;  
 Best Local Similarity 100.0%; Pred. No. 0.00094;  
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KLVPPFATELHERL 13  
 DB 79 KLVPPFATELHERL 91  
 |||||

RESULT 42  
 AAU10861  
 ID AAU10861 standard; protein; 396 AA.

XX AC

XX AC AAU10861;

XX DT 14-FEB-2002 (first entry)

XX DE Human apolipoprotein A-IV, APOA4, isoform #1.

XX KW Human; APOA4; apolipoprotein A-IV; antiatherosclerotic; cardiant;  
 KW haplotype; chromosome 11q23-qter; coronary heart disease; obesity;  
 KW atherosclerosis.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

XX FT Misc-difference 13 /note= "Wild-type Val substituted by Met"

XX FT WO20017124-A2.

XX PD 18-OCT-2001.

XX PF 03-APR-2001; 2001WO-US010670.

XX PR 05-APR-2000; 2000US-0194362P.

XX PA (GENA-) GENAISSANCE PHARM INC.

XX PI Bentivegna SC, Choi JY, Kliem SE, Koshy B;

XX DR WPI; 2002-041281/05.

XX PT New haplotypes of the human apolipoprotein A-IV gene, useful to diagnose  
 PT and treat disorders associated with its abnormal expression or function  
 PT such as coronary artery disease.

XX PS Claim 29; Page; 71pp; English.

XX CC The invention relates to haplotyping the human apolipoprotein A-IV  
 CC (APOA4) gene of an individual, comprising determining if the individual  
 CC has one of the APOA4 haplotypes or haplotype pairs fully defined in the  
 CC specification. Also disclosed are genotyping oligonucleotides (or allele  
 CC specific oligonucleotides, ASO) as well as methods for correlating a  
 CC particular haplotype pair with a trait e.g. obesity, in a population. The  
 CC APOA4 gene is located on chromosome 11q23-qter. The methods of the

CC invention are useful to diagnose and develop treatment for disorders  
 CC associated with abnormal APOA4 expression or function, for example  
 CC coronary heart disease and atherosclerosis. The APOA4 isogenes and  
 CC screened compounds are useful for the treatment of disorders associated  
 CC with abnormal APOA4 expression or function such as coronary artery  
 CC disease. The present sequence is an isoform of the APOA4 protein. Note:  
 CC The present sequence is not shown in the specification but was created by  
 CC the indexer from the APOA4 sequence shown in figure 3

XX SQ Sequence 396 AA;

Query Match 100.0%; Score 66; DB 5; Length 396;  
 Best Local Similarity 100.0%; Pred. No. 0.00094;  
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KLVPPFATELHERL 13  
 DB 79 KLVPPFATELHERL 91  
 |||||

RESULT 43

AAU10868  
 ID AAU10868 standard; protein; 396 AA.

XX AC

XX AC AAU10868;

XX DT 14-FEB-2002 (first entry)

XX DE Human apolipoprotein A-IV, APOA4, isoform #8.

XX KW Human; APOA4; apolipoprotein A-IV; antiatherosclerotic; cardiant;  
 KW haplotype; chromosome 11q23-qter; coronary heart disease; obesity;  
 KW atherosclerosis.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

XX FT Misc-difference 367 /note= "Wild-type Thr substituted by Ser"

XX FT WO20017124-A2.

XX PD 18-OCT-2001.

XX PF 03-APR-2001; 2001WO-US010670.

XX PR 05-APR-2000; 2000US-0194362P.

XX PA (GENA-) GENAISSANCE PHARM INC.

XX PI Bentivegna SC, Choi JY, Kliem SE, Koshy B;

XX DR WPI; 2002-041281/05.

XX PT New haplotypes of the human apolipoprotein A-IV gene, useful to diagnose  
 PT and treat disorders associated with its abnormal expression or function  
 PT such as coronary artery disease.

XX PS Claim 29; Page; 71pp; English.

XX CC The invention relates to haplotyping the human apolipoprotein A-IV  
 CC (APOA4) gene of an individual, comprising determining if the individual  
 CC has one of the APOA4 haplotypes or haplotype pairs fully defined in the  
 CC specification. Also disclosed are genotyping oligonucleotides (or allele  
 CC specific oligonucleotides, ASO) as well as methods for correlating a  
 CC particular haplotype pair with a trait e.g. obesity, in a population. The  
 CC APOA4 gene is located on chromosome 11q23-qter. The methods of the  
 CC invention are useful to diagnose and develop treatment for disorders  
 CC associated with abnormal APOA4 expression or function, for example  
 CC coronary heart disease and atherosclerosis. The APOA4 isogenes and  
 CC screened compounds are useful for the treatment of disorders associated  
 CC with abnormal APOA4 expression or function such as coronary artery  
 CC disease. The present sequence is an isoform of the APOA4 protein. Note:

CC The present sequence is not shown in the specification but was created by  
 CC the indexer from the APOA4 sequence shown in figure 3

XX SQ Sequence 396 AA;

Query Match 100.0%; Score 66; DB 5; Length 396;  
 Best Local Similarity 100.0%; Pred. No. 0.00094;  
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLVPFATLHERL 13  
 |||||

Db 79 KLVPFATLHERL 91

RESULT 44  
 AAU10860  
 ID AAU10860 standard; protein; 396 AA.

XX AC AAU10860;

DT 14-FEB-2002 (first entry)

XX DE Human apolipoprotein A-IV, APOA4.

XX KW Human; APOA4; apolipoprotein A-IV; antiatherosclerotic; cardiant;  
 KW haplotype; chromosome 11q23-qter; coronary heart disease; obesity;  
 KW atherosclerosis.

XX OS Homo sapiens.

XX PN WO200177124-A2.

XX PD 18-OCT-2001.

XX PF 03-APR-2001; 2001WO-US010670.

XX PR 05-APR-2000; 2000US-0194362P.

XX PA (GENA-) GENAISSANCE PHARM INC.

XX PI Bentivegna SC, Choi JY, Kliem SE, Koshy B;

XX WPI; 2002-041281/05.

XX DR N-PSDB; AAS16512, AAS16513.

XX New haplotypes of the human apolipoprotein A-IV gene, useful to diagnose  
 PT and treat disorders associated with its abnormal expression or function  
 PT such as coronary artery disease.

XX PS Claim 29; Fig 3; 71pp; English.

XX The invention relates to haplotyping the human apolipoprotein A-IV  
 CC (APOA4) gene of an individual, comprising determining if the individual  
 CC has one of the APOA4 haplotypes or haplotype pairs fully defined in the  
 CC specification. Also disclosed are genotyping oligonucleotides (or allele  
 CC specific oligonucleotides, ASO) as well as methods for correlating a  
 CC particular haplotype pair with a trait e.g. obesity, in a population. The  
 CC APOA4 gene is located on chromosome 11q23-qter. The methods of the  
 CC invention are useful to diagnose and develop treatment for disorders  
 CC associated with abnormal APOA4 expression or function, for example  
 CC coronary heart disease and atherosclerosis. The APOA4 isogenes and  
 CC screened compounds are useful for the treatment of disorders associated  
 CC with abnormal APOA4 expression or function such as coronary artery  
 CC disease. The present sequence is the APOA4 protein

XX SQ Sequence 396 AA;

Query Match 100.0%; Score 66; DB 5; Length 396;  
 Best Local Similarity 100.0%; Pred. No. 0.00094;  
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLVPFATLHERL 13  
 |||||

Db 79 KLVPFATLHERL 91

RESULT 45

AAU10864  
 ID AAU10864 standard; protein; 396 AA.

XX AC AAU10864;

DT 14-FEB-2002 (first entry)

XX DE Human apolipoprotein A-IV, APOA4, isoform #4.

XX KW Human; APOA4; apolipoprotein A-IV; antiatherosclerotic; cardiant;  
 KW haplotype; chromosome 11q23-qter; coronary heart disease; obesity;  
 KW atherosclerosis.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers  
 FT Misc-difference 147  
 FT /note= "Wild-type Asn substituted by Ser"

XX PN WO200177124-A2.

XX PD 18-OCT-2001.

XX PF 03-APR-2001; 2001WO-US010670.

XX PR 05-APR-2000; 2000US-0194362P.

XX PA (GENA-) GENAISSANCE PHARM INC.

XX PI Bentivegna SC, Choi JY, Kliem SE, Koshy B;

XX WPI; 2002-041281/05.

XX New haplotypes of the human apolipoprotein A-IV gene, useful to diagnose  
 PT and treat disorders associated with its abnormal expression or function  
 PT such as coronary artery disease.

XX PS Claim 29; Page; 71pp; English.

XX The invention relates to haplotyping the human apolipoprotein A-IV  
 CC (APOA4) gene of an individual, comprising determining if the individual  
 CC has one of the APOA4 haplotypes or haplotype pairs fully defined in the  
 CC specification. Also disclosed are genotyping oligonucleotides (or allele  
 CC specific oligonucleotides, ASO) as well as methods for correlating a  
 CC particular haplotype pair with a trait e.g. obesity, in a population. The  
 CC APOA4 gene is located on chromosome 11q23-qter. The methods of the  
 CC invention are useful to diagnose and develop treatment for disorders  
 CC associated with abnormal APOA4 expression or function, for example  
 CC coronary heart disease and atherosclerosis. The APOA4 isogenes and  
 CC screened compounds are useful for the treatment of disorders associated  
 CC with abnormal APOA4 expression or function such as coronary artery  
 CC disease. The present sequence is an isoform of the APOA4 protein. Note:  
 CC The present sequence is not shown in the specification but was created by  
 CC the indexer from the APOA4 sequence shown in figure 3

XX SQ Sequence 396 AA;

Query Match 100.0%; Score 66; DB 5; Length 396;  
 Best Local Similarity 100.0%; Pred. No. 0.00094;  
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLVPFATLHERL 13  
 |||||

Db 79 KLVPFATLHERL 91

Search completed: September 1, 2005, 16:37:42  
 Job time : 170 secs

**This Page Blank (uspto)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 1, 2005, 16:21:35 ; Search time 174 Seconds  
(without alignments)  
38.259 Million cell updates/sec

Title: US-09-993-366-1

Perfect score: 66

Sequence: 1 KLVPFATLHERL 13

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database :

UniProt\_03.\*

1: uniprot\_sprot.\*

2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	66	100.0	102	2 Q29258	Q29258 sus scrofa
2	66	100.0	382	1 APA4_PIG	Q46409 sus scrofa
3	66	100.0	386	1 APA4_HUMAN	P06727 homo sapien
4	66	100.0	401	1 APA4_PAPAN	Q28758 papio anubi
5	66	100.0	429	1 APA4_MACFA	F33621 macaca fasc
6	45	68.2	366	2 Q93601	Q93601 gallus gall
7	42	63.6	436	1 DADA_PHOLL	Q7n326 photorhabdu
8	42	63.6	4345	2 Q9VLA0	Q9VLA0 drosophila
9	42	63.6	4374	2 Q7K7G2	Q7K7G2 drosophila
10	41	62.1	258	2 Q08877	Q08877 rattus norv
11	41	62.1	258	2 Q09054	Q09054 rattus norv
12	41	62.1	261	2 Q8U799	Q8U799 agrobacteri
13	41	62.1	265	2 Q7CW91	Q7cw91 agrobacteri
14	41	62.1	367	2 Q6P7H6	Q6p7h6 xenopus lae
15	41	62.1	897	2 Q89EW7	Q89ew7 bradyrhizob
16	41	62.1	1056	2 Q7Q9K3	Q7q9k3 corynebacte
17	41	62.1	1179	2 Q6NU57	Q6nj57 corynebacte
18	40	60.6	260	1 TNF5_RAT	Q922v2 rattus norv
19	40	60.6	266	1 APA1_RABIT	P09809 cryctolagus
20	40	60.6	368	2 Q66IL6	Q66il6 xenopus tro
21	40	60.6	585	2 Q6CGG7	Q6cgs7 yarrowia li
22	40	60.6	656	2 Q7ULM6	Q7ulm6 rhodospirell
23	39	59.1	150	2 Q7N819	Q7n819 photorhabdu
24	39	59.1	177	2 Q8GF72	Q8gf72 photorhabdu
25	39	59.1	220	2 Q830L3	Q830l3 enterococcu
26	39	59.1	353	2 P87210	P87210 neurospora
27	39	59.1	697	2 Q9CAI9	Q9caia9 arabidopsis
28	39	59.1	2710	2 Q7R3A9	Q7r3a9 giardia lam
29	39	59.1	2796	2 Q48926	Q48926 mycobacteri
30	39	59.1	3069	2 Q7D623	Q7d623 mycobacteri
31	39	59.1	3069	2 P95029	P95029 mycobacteri

#### ALIGNMENTS

32	39	59.1	3069	2	Q7TYD8	Q7tyd8 mycobacteri
33	39	59.1	3076	2	Q9X7E2	Q9x7e2 mycobacteri
34	39	59.1	3092	2	Q73XH7	Q73xh7 mycobacteri
35	38	57.6	57	2	Q6K305	Q6k305 oryza sativ
36	38	57.6	162	2	Q89SR9	Q89sr9 bradyrhizob
37	38	57.6	176	2	Q6HCP9	Q6hcp9 bacillus th
38	38	57.6	250	2	Q9X548	Q9x548 corynebacte
39	38	57.6	288	2	Q33624	Q33624 streptomyce
40	38	57.6	311	2	Q6GK71	Q6gk71 staphylococ
41	38	57.6	314	2	Q6RWS4	Q6rws4 uncultured
42	38	57.6	317	2	Q6G538	Q6g538 aquifex aeo
43	38	57.6	319	2	Q8TYA1	Q8tya1 methanopyru
44	38	57.6	330	2	Q67L63	Q67l63 symbiobacte
45	38	57.6	335	2	Q6R3I5	Q6r3i5 gemmata sp.
46	38	57.6	354	2	Q8FC11	Q8fc11 escherichia
47	38	57.6	391	1	APA4_RAT	P02651 rattus norv
48	38	57.6	394	2	Q89IC2	Q89ic2 clostridium
49	38	57.6	405	1	PURA_DEIRA	Q9ryb5 deinococcus
50	38	57.6	423	2	Q75TW9	Q75tw9 bacillus ha
51	38	57.6	428	1	PURA_BACHD	Q9k5r0 bacillus ha
52	38	57.6	428	1	PURA_BACAN	Q81j19 bacillus an
53	38	57.6	429	1	PURA_BACCR	Q814h1 bacillus ce
54	38	57.6	429	1	PURA_OCEIH	Q8ekx9 oceanobacil
55	38	57.6	429	2	Q630D5	Q630d5 bacillus ce
56	38	57.6	429	2	Q72MW1	Q72mw1 bacillus ce
57	38	57.6	429	2	Q6HAG9	Q6hag9 bacillus th
58	38	57.6	432	1	PURA_AGR75	Q8ucn6 agrobacteri
59	38	57.6	472	1	Q7SA_SALTY	Q91893 salmonella
60	38	57.6	474	2	Q7BJX2	Q7bjx2 plasmodium
61	38	57.6	543	2	Q28264	Q28264 archaeoglob
62	38	57.6	589	2	Q83989	Q83989 treponema p
63	38	57.6	652	2	Q82C93	Q82c93 streptomyce
64	38	57.6	661	2	Q9L1H1	Q9l1h1 streptomyce
65	38	57.6	669	1	HPO_DROME	Q8t066 drosophila
66	38	57.6	711	2	Q8D4D6	Q8d4d6 vibrio vuln
67	38	57.6	715	2	Q7MFW4	Q7mfw4 vibrio vuln
68	38	57.6	717	2	Q74Z02	Q74z02 ashbya goss
69	38	57.6	847	1	Q9SLU4	Q9slu4 arabidopsis
70	38	57.6	849	1	RS62_HUMAN	Q1283 homo sapien
71	38	57.6	864	2	Q6YQ88	Q6yqr8 onion yello
72	38	57.6	983	1	GCSP_ANASP	Q6ynf9 anabaena sp
73	38	57.6	987	1	TNPA_BACTU	P10021 bacillus th
74	38	57.6	1330	2	Q9SXQ2	Q9sxq2 arabidopsis
75	38	57.6	1421	2	Q9SXQ5	Q9sxq5 arabidopsis
76	38	57.6	1431	2	Q9SXQ1	Q9sxq1 arabidopsis
77	38	57.6	1447	2	Q9SXQ3	Q9sxq3 arabidopsis
78	38	57.6	1456	2	Q9ZT94	Q9zt94 arabidopsis
79	38	57.6	1459	2	Q94LA8	Q94la8 arabidopsis
80	38	57.6	1466	2	Q94HM2	Q94hw2 arabidopsis
81	38	57.6	1466	2	Q94HW7	Q94hw7 arabidopsis
82	38	57.6	1474	2	Q23529	Q23529 arabidopsis
83	38	57.6	1475	2	Q9SXQ4	Q9sxq4 arabidopsis
84	38	57.6	1592	2	Q8U110	Q8j110 cryptococcu
85	38	57.6	1593	2	Q8J0X4	Q8j0x4 cryptococcu
86	38	57.6	4600	2	Q8IFF9	Q8iff9 trypanosoma
87	37	56.1	95	2	Q7PLK4	Q7plk4 drosophila
88	37	56.1	100	2	Q50146	Q50146 mycobacteri
89	37	56.1	147	2	Q8X1J3	Q8x1j3 clostridium
90	37	56.1	176	2	Q63LD3	Q63ld3 burkholderi
91	37	56.1	209	2	Q8GMR9	Q8gmr9 synecococc
92	37	56.1	222	2	Q9XXI7	Q9xxi7 caenorhabdi
93	37	56.1	226	2	Q7Q041	P04639 rattus norv
94	37	56.1	259	1	APA1_RAT	Q70cc4 eimeria ten
95	37	56.1	263	2	Q70CC4	P18648 sus scrofa
96	37	56.1	265	1	APA1_PIG	Q70cd3 eimeria ten
97	37	56.1	268	2	Q70CD2	Q70cd3 eimeria ten
98	37	56.1	270	2	Q70CD3	Q72tt1 leptospira
99	37	56.1	279	2	Q72TT1	Q8f1e4 leptospira
100	37	56.1	279	2	Q8F1E4	

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

-----

CC EMBL; AJ222966; CAA11020.1; -.

DR HSSP; P32851; 1E23.

DR InterPro; IPR009074; Apolipo\_A\_E\_C3.

DR InterPro; IPR000074; Apolipoprotein.

DR Pfam; PF01442; Apolipoprotein; 2.

KW Chylomicron; HDL; Lipid transport; Plasma; Repeat; Signal.

FT SIGNAL 1 20

FT CHAIN 21 382

FT DOMAIN 33 330

FT REPEAT 33 54

FT REPEAT 60 81

FT REPEAT 82 103

FT REPEAT 115 136

FT REPEAT 137 158

FT REPEAT 159 180

FT REPEAT 181 202

FT REPEAT 203 224

FT REPEAT 225 246

FT REPEAT 247 268

FT REPEAT 269 286

FT REPEAT 287 308

FT REPEAT 309 330

FT DOMAIN 360 377

SQ SEQUENCE 382 AA; 43294 MW; 2ACA88A02D4379BF CRC64;

Query Match 100.0%; Score 66; DB 1; Length 382;

Best Local Similarity 100.0%; Pred. No. 0.0025;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KLVPPFATLHERL 13

DB 79 KLVPPFATLHERL 91

RESULT 3

AP04 HUMAN

ID AP04 HUMAN STANDARD; PRT; 396 AA.

AC P06727;

DT 01-JAN-1988 (Rel. 06, Created)

DT 01-FEB-1991 (Rel. 17, Last sequence update)

DT 05-JUL-2004 (Rel. 44, Last annotation update)

DE Apolipoprotein A-IV precursor (Apo-AIV).

GN Name=APOA4;

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI\_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Intestine;

RX MEDLINE=89194198; PubMed=2930771; DOI=10.1016/0005-2760(89)90292-0;

RA Gatto C., Gu Z.W., Xiong W., Rosseneu M., Yang H.X., Lee B.M.,

RA Gotto A.M. Jr., Chan L.

RT "The primary structure of human apolipoprotein A-IV."

RL Biochim. Biophys. Acta 1002:231-237(1988).

RN [2]

RP SEQUENCE FROM N.A.

RX MEDLINE=87041474; PubMed=3095836;

RA Karathanasis S.K., Oettingen P., Haddad I.A., Antonarakis S.E.;

RT "Structure, evolution, and polymorphisms of the human apolipoprotein A4 gene (APOA4)."

RL Proc. Natl. Acad. Sci. U.S.A. 83:8457-8461(1986).

RN [3]

RP SEQUENCE FROM N.A.

RX MEDLINE=86296629; PubMed=3755616;

QY 1 KLVPPFATLHERL 13

DB 79 KLVPPFATLHERL 91

RESULT 2

AP04 FIG

ID AP04 FIG STANDARD; PRT; 382 AA.

AC O46409;

DT 15-DEC-1998 (Rel. 37, Created)

DT 15-DEC-1998 (Rel. 37, Last sequence update)

DT 05-JUL-2004 (Rel. 44, Last annotation update)

DE Apolipoprotein A-IV precursor (Apo-AIV).

GN Name=APOA4;

OS Sus scrofa (Pig).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.

OX NCBI\_TaxID=9823;

RN [1]

RP SEQUENCE FROM N.A.

RX Oasda J., Iturralde M., Calleja L., Gonzalez N., Pineiro A.;

RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.

CC -I- FUNCTION: May have a role in chylomicrons and VLDL secretion and catabolism. Required for efficient activation of lipoprotein lipase by ApoC-II; potent activator of LCAT. ApoA-IV is a major component of HDL and chylomicrons.

CC -I- SUBCELLULAR LOCATION: Secreted.

CC -I- TISSUE SPECIFICITY: Secreted in plasma.

CC -I- DOMAIN: Nine of the thirteen 22-amino acid tandem repeats (each 22-mer is actually a tandem array of two, A and B, related 11-mers) occurring in this sequence are predicted to be highly alpha-helical, and many of these helices are amphipathic. They may therefore serve as lipid-binding domains with lecithin:cholesterol acyltransferase (LCAT) activating abilities.

CC -I- SIMILARITY: Belongs to the apolipoprotein A1/A4/E family.

RA Karathanasis S.K., Yunis I.;  
 RT "Structure, evolution, and tissue-specific synthesis of human  
 RT apolipoprotein AIV.";  
 RL Biochemistry 25:3962-3970(1986).  
 [4]  
 RN SEQUENCE FROM N.A.  
 RP MEDLINE=87250378; PubMed=3036793;  
 RX Elshourbagy N.A., Walker D.W., Paik Y.K., Boguski M.S., Freeman M.,  
 RA Gordon J.I., Taylor J.M.;  
 RT "Structure and expression of the human apolipoprotein A-IV gene.";  
 RL J. Biol. Chem. 262:7973-7981(1987).  
 [5]  
 RN SEQUENCE OF 21-396 FROM N.A.  
 RP MEDLINE=86111895; PubMed=3080432;  
 RX Elshourbagy N.A., Walker D.W., Boguski M.S., Gordon J.I., Taylor J.M.;  
 RA "The nucleotide and derived amino acid sequence of human  
 RT apolipoprotein A-IV mRNA and the close linkage of its gene to the  
 RT genes of apolipoproteins A-I and C-III.";  
 RL J. Biol. Chem. 261:1998-2002(1986).  
 [6]  
 RN SIGNAL SEQUENCE CLEAVAGE SITE.  
 RP MEDLINE=84161950; PubMed=6706947;  
 RX Gordon J.I., Bisgaler C.L., Sims H.F., Sachdev O.P., Glickman R.M.,  
 RA Strauss A.W.;  
 RT "Biosynthesis of human preapolipoprotein A-IV.";  
 RL J. Biol. Chem. 259:468-474(1984).  
 [7]  
 RN REVIEW ON POLYMORPHISM.  
 RP Lohse P., Brewer H.B. Jr.;  
 RA "Genetic polymorphism of apolipoprotein A-IV.";  
 RL Curr. Opin. Lipidol. 2:90-95(1991).  
 [8]  
 RN VARIANT A-IV\*2.  
 RP MEDLINE=90277616; PubMed=2351649;  
 RX Lohse P., Kindt M.R., Rader D.J., Brewer H.B. Jr.;  
 RA "Genetic polymorphism of human plasma apolipoprotein A-IV is due to  
 RT nucleotide substitutions in the apolipoprotein A-IV gene.";  
 RL J. Biol. Chem. 265:10061-10064(1990).  
 [9]  
 RN VARIANTS A-IV\*0 AND A-IV\*3.  
 RP MEDLINE=90324273; PubMed=1973689;  
 RX Lohse P., Kindt M.R., Rader D.J., Brewer H.B. Jr.;  
 RA "Human plasma apolipoproteins A-IV\*0 and A-IV\*3. Molecular basis for  
 RT two rare variants of apolipoprotein A-IV\*1.";  
 RL J. Biol. Chem. 265:12734-12739(1990).  
 [10]  
 RN VARIANTS.  
 RP MEDLINE=91310615; PubMed=1677358;  
 RX Lohse P., Kindt M.R., Rader D.J., Brewer H.B. Jr.;  
 RA "Three genetic variants of human plasma apolipoprotein A-IV: apoA-IV-  
 RT 1(Thr-347-->Ser), apoA-IV\*0(Lys-167-->Glu,Gln-360-->His), and apoA-IV-  
 RT 3(Glu-165-->Lys).";  
 RL J. Biol. Chem. 266:13513-13518(1991).  
 [11]  
 RN ERRATUM.  
 RP Lohse P., Kindt M.R., Rader D.J., Brewer H.B. Jr.;  
 RA J. Biol. Chem. 266:19866-19866(1991).  
 [12]  
 RN VARIANT MET-13.  
 RP MEDLINE=92238494; PubMed=1349197;  
 RX von Eckardstein A., Funke H., Schulte M., Erren M., Schulte H.,  
 RA Asmann G.;  
 RT "Nonsynonymous polymorphic sites in the apolipoprotein (apo) A-IV gene  
 RT are associated with changes in the concentration of apo B- and apo A-  
 RT I-containing lipoproteins in a normal population.";  
 RL Am. J. Hum. Genet. 50:1115-1128(1992).  
 [13]  
 RN VARIANT SER-147.  
 RP MEDLINE=92144647; PubMed=1737067; DOI=10.1016/0925-4439(92)90147-F;  
 RX Tenkanen H., Koskinen P., Metsä J., Baumann M., Lukka M.,  
 RA Kauppinen-Makelin R., Kontula K., Taskinen M.R., Manttari M.,  
 RA Manninen V., Ehnholm C.;  
 RT "A novel polymorphism of apolipoprotein A-IV is the result of an

RT asparagine to serine substitution at residue 127.";  
 RL Biochim. Biophys. Acta 1138:27-33(1992).  
 [14]  
 RN VARIANT A-IV\*5.  
 RP MEDLINE=93138374; PubMed=14871136;  
 RX Kamboh M.I., Williams E.R., Law J.C., Aston C.E., Bunker C.H.,  
 RA Ferrell R.E., Pollitzer W.S.;  
 RT "Molecular basis of a unique African variant (A-IV 5) of human  
 RT apolipoprotein A-IV and its significance in lipid metabolism.";  
 RL Genet. Epidemiol. 9:379-388(1992).  
 [15]  
 RN VARIANTS BUDAPEST-2 LYS-44; BUDAPEST-1 CYS-305 AND SER-367.  
 RP MEDLINE=95245341; PubMed=7728150;  
 RX Menzel H.J., Dieplinger H., Sandholzer C., Karadi I., Utermann G.,  
 RA Csaszar A.;  
 RT "Apolipoprotein A-IV polymorphism in the Hungarian population: Gene  
 RT frequencies, effect on lipid levels, and sequence of two new  
 RT variants.";  
 RL Hum. Mutat. 5:58-65(1995).  
 [16]  
 RN VARIANTS SEATTLE-3 SER-161; SEATTLE-1 LEU-178 AND SEATTLE-2 GLN-264.  
 RP MEDLINE=97114287; PubMed=8956036;  
 RX DOI=10.1002/(SICI)1098-1004(1996)8:4<319::AID-HUMU4>3.3.CO;2-T;  
 RA Deeb S.S., Nevin D.N., Iwasaki L., Brunzell J.D.;  
 RT "Two novel apolipoprotein A-IV variants in individuals with familial  
 RT combined hyperlipidemia and diminished levels of lipoprotein lipase  
 RT activity.";  
 RL Hum. Mutat. 8:319-325(1996).  
 [17]  
 RN VARIANT HIS-380.  
 RP MEDLINE=99318094; PubMed=10391210; DOI=10.1038/10297;  
 RX Halushka M.K., Fan J.-B., Bentley K., Hsie L., Shen N., Weder A.,  
 RA Cooper R., Lipshutz R., Chakravarti A.;  
 RT "Patterns of single-nucleotide polymorphisms in candidate genes for  
 RT blood-pressure homeostasis.";  
 RL Nat. Genet. 22:239-247(1999).  
 CC -I- FUNCTION: May have a role in chylomicrons and VLDL secretion and  
 CC catabolism. Required for efficient activation of lipoprotein  
 CC lipase by ApoC-II; potent activator of LCAT. ApoA-IV is a major  
 CC component of HDL and chylomicrons.  
 CC -I- SUBCELLULAR LOCATION: Secreted.  
 CC -I- TISSUE SPECIFICITY: Synthesized primarily in the intestine and  
 CC secreted in plasma.  
 CC -I- DOMAIN: Nine of the thirteen 22-amino acid tandem repeats (each  
 CC 22-mer is actually a tandem array of two, A and B, related 11-  
 CC mers) occurring in this sequence are predicted to be highly alpha-  
 CC helical, and many of these helices are amphipathic. They may  
 CC therefore serve as lipid-binding domains with lecithin:cholesterol  
 CC acyltransferase (LCAT) activating abilities.  
 CC -I- POLYMORPHISM: Eight alleles have been characterized (APOA-IV\*0 to  
 CC APOA-IV\*7). APOA-IV\*1 is the major allele (90%), APOA-IV\*2 is also  
 CC common (8%), the others are rare alleles.  
 CC -I- SIMILARITY: Belongs to the apolipoprotein A1/A4/E family.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 CC or send an email to license@isb-sib.ch).  
 CC  
 CC EMBL; M14642; AAA51745.1; -;  
 CC EMBL; X13629; CAA31955.1; -;  
 CC EMBL; M14566; AAA51748.1; -;  
 CC EMBL; J02758; AAA96731.1; -;  
 CC EMBL; M13654; AAA51744.1; -;  
 CC PIR; A94137; LPHU44.  
 CC HSSP; P02649; IOR2.  
 CC SWISS-2DPAGE; P06727; HUMAN.  
 CC Genew; HGNC:602; APOA4.  
 CC MIM; 107690; -;  
 CC GO; GO:0005319; F:lipid transporter activity; TAS.

```

DR GO:0008015; P:circulation; TAS.
DR GO:0006869; P:lipid transport; TAS.
DR InterPro: IPR009074; Apolipo A_E C3.
DR InterPro: IPR000074; Apolipoprotein.
DR Pfam: PF01442; Apolipoprotein; 2.
KW Chylomicron; HDL; Lipid transport; Plasma; Polymorphism; Repeat;
Signal.
FT SIGNAL 1 20 Apolipoprotein A-IV.
FT CHAIN 21 396 13 X 22 AA approximate tandem repeats.
FT DOMAIN 33 330 1.
FT REPEAT 33 54 2.
FT REPEAT 60 81 3.
FT REPEAT 82 103 4.
FT REPEAT 115 136 5.
FT REPEAT 137 158 6.
FT REPEAT 159 180 7.
FT REPEAT 181 202 8.
FT REPEAT 203 224 9.
FT REPEAT 225 246 10.
FT REPEAT 247 268 11.
FT REPEAT 269 286 12.
FT REPEAT 287 308 13.
FT REPEAT 309 330 Gln/Glu-rich.
FT DOMAIN 372 389 V -> M (in allele APOA-IV*1D).
FT VARIANT 13 13

Query Match 100.0%; Score 66; DB 1; Length 396;
Best Local Similarity 100.0%; Pred. No. 0.0025;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KLVPPFATLHERL 13
Db 79 KLVPPFATLHERL 91

RESULT 4
ID AP44 PAPAN STANDARD; PRT; 401 AA.
AC Q28756;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Apolipoprotein A-IV precursor (Apo-AIV) (Fragment).
GN Name=APOA4;
OS Papio anubis (Olive baboon).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecoidea; Papio.
OX NCBI_TaxID=9555;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Intestine;
RX MEDLINE=93340170; PubMed=8101842;
RA Hixson J.E., Kammerer C.M., Mott G.E., Britten M.L., Birnbaum S.,
RA Powers P.K., Vandenberg J.L.;
RT "Baboon apolipoprotein A-IV. Identification of Lys-76-->Glu that
RT distinguishes two common isoforms and detection of length
RT polymorphisms at the carboxyl terminus.";
RL J. Biol. Chem. 268:15667-15673(1993).
CC -1- FUNCTION: May have a role in chylomicrons and VLDL secretion and
CC catabolism. Required for efficient activation of lipoprotein
CC lipase by ApoC-II; potent activator of LCAT. ApoA-IV is a major
CC component of HDL and chylomicrons.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- TISSUE SPECIFICITY: Synthesized primarily in the intestine and
CC secreted in plasma.
CC -1- DOMAIN: Nine of the thirteen 22-amino acid tandem repeats (each
CC 22-mer is actually a tandem array of two, A and B, related 11-
CC mers) occurring in this sequence are predicted to be highly alpha-
CC helical, and many of these helices are amphipathic. They may
CC therefore serve as lipid-binding domains with lecithin:cholesterol
CC acyltransferase (LCAT) activating abilities.
CC -1- POLYMORPHISM: There are two common APOA-IV isoforms, I (shown

```

```

CC here) and E. The I isoform is associated with higher levels of
CC high density lipoprotein-C on a high cholesterol, saturated fat
CC diet.
CC -1- SIMILARITY: Belongs to the apolipoprotein A1/A4/E family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; L13174; AAA35379.1; -.
DR HSSP; P02649; 1B68.
DR InterPro: IPR009074; Apolipo A_E C3.
DR InterPro: IPR000074; Apolipoprotein.
DR Pfam: PF01442; Apolipoprotein; 2.
KW Chylomicron; HDL; Lipid transport; Plasma; Polymorphism; Repeat;
Signal.
FT SIGNAL 1 1 Potential.
FT CHAIN <1 4 Apolipoprotein A-IV.
FT DOMAIN 17 314 13 X 22 AA approximate tandem repeats.
FT REPEAT 17 38 1.
FT REPEAT 44 65 2.
FT REPEAT 66 87 3.
FT REPEAT 99 120 4.
FT REPEAT 121 142 5.
FT REPEAT 143 164 6.
FT REPEAT 165 186 7.
FT REPEAT 187 208 8.
FT REPEAT 209 230 9.
FT REPEAT 231 252 10.
FT REPEAT 253 270 11.
FT REPEAT 271 292 12.
FT REPEAT 293 314 13.
FT DOMAIN 356 394 Gln/Glu-rich.
FT VARIANT 80 80 K -> E (in isoform E).
SQ SEQUENCE 401 AA; 46538 MW; 0A76D1284AA9837F CRC64;

Query Match 100.0%; Score 66; DB 1; Length 401;
Best Local Similarity 100.0%; Pred. No. 0.0026;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KLVPPFATLHERL 13
Db 63 KLVPPFATLHERL 75

RESULT 5
ID AP44 MACFA STANDARD; PRT; 429 AA.
AC P33621;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Apolipoprotein A-IV precursor (Apo-AIV).
GN Name=APOA4;
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecoidea; Macaca.
OX NCBI_TaxID=9541;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Leukocyte;
RX MEDLINE=93192330; PubMed=8448212; DOI=10.1016/0167-4781(93)90226-4;
RA Osada J., Pocovi M., Nicolosi R.J., Schaefer E.J., Ordovas J.M.;
RT "Nucleotide sequences of the Macaca fascicularis apolipoprotein C-III
RT and A-IV genes.";
RL Biochim. Biophys. Acta 1172:335-339(1993).
CC -1- FUNCTION: May have a role in chylomicrons and VLDL secretion and

```

catabolism. Required for efficient activation of lipoprotein lipase by ApoC-II; potent activator of LCAT. ApoA-IV is a major component of HDL and chylomicrons.

CC -1- SUBCELLULAR LOCATION: Secreted.

CC -1- TISSUE SPECIFICITY: Secreted in plasma.

CC -1- DOMAIN: Nine of the thirteen 22-amino acid tandem repeats (each 22-mer is actually a tandem array of two, A and B, related 11-mers) occurring in this sequence are predicted to be highly alpha-helical, and many of these helices are amphipathic. They may therefore serve as lipid-binding domains with lecithin:cholesterol acyltransferase (LCAT) activating abilities.

CC -1- SIMILARITY: Belongs to the apolipoprotein AI/A4/E family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

CC EMBL; X68361; CAA48421.1; --

CC PIR; S30195; S29565.

CC HSSP; P02649; 1B68.

CC InterPro; IPR009074; Apolipo A E C3.

CC InterPro; IPR000074; Apolipoprotein.

CC Pfam; PF01442; Apolipoprotein; 2.

CC Chylomicron; HDL; Lipid transport; Plasma; Repeat; Signal.

CC SIGNAL 1 20 By similarity.

CC CHAIN 21 429 Apolipoprotein A-IV.

CC DOMAIN 33 330 13 X 22 AA approximate tandem repeats.

CC REPEAT 33 330 1.

CC REPEAT 60 81 2.

CC REPEAT 82 103 3.

CC REPEAT 115 136 4.

CC REPEAT 137 158 5.

CC REPEAT 159 180 6.

CC REPEAT 181 202 7.

CC REPEAT 203 224 8.

CC REPEAT 225 246 9.

CC REPEAT 247 268 10.

CC REPEAT 269 286 11.

CC REPEAT 287 308 12.

CC REPEAT 309 330 13.

CC DOMAIN 372 420 Gln/Glu-rich.

CC SEQUENCE 429 AA; 49876 MW; 3D458F551D0DB60C CRC64;

Query Match 100.0%; Score 66; DB 1; Length 429;

Best Local Similarity 100.0%; Pred. No. 0.0028;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KLVPFATLHERL 13

DB 79 KLVPFATLHERL 91

RESULT 6

ID O93601 PRELIMINARY; PRT; 366 AA.

AC O93601;

DT 01-NOV-1998 (TrEMBLrel. 08, Created)

DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)

DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)

DE Apolipoprotein AIV.

GN Name=apoAIV;

OS Gallus gallus (Chicken).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;

OC Gallus.

OX NCBI\_TaxID=9031;

OX [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=98221191; PubMed=9553114; DOI=10.1074/jbc.273.17.10543;

RA Steinmetz A., Hermann M., Nimpf J., Abersold R., Ducret A., Weinberg R.B., Schneider W.J.;

RT "Expression and conservation of apolipoprotein AIV in an avian species";

RL J. Biol. Chem. 273:10543-10549(1998).

DR EMBL; Y16534; CAA76273.1; --

GO GO:0005576; C:extracellular; IEA.

GO GO:0008289; F:lipid binding; IEA.

GO GO:0006869; P:lipid transport; IEA.

GO GO:0042157; P:lipoprotein metabolism; IEA.

DR InterPro; IPR000074; Apolipoprotein.

DR InterPro; IPR009074; Apolipo A E C3.

DR Pfam; PF01442; Apolipoprotein; 2.

KW Lipoprotein.

SQ SEQUENCE 366 AA; 40853 MW; B47DB49F3E8D91CE CRC64;

Query Match 68.2%; Score 45; DB 2; Length 366;

Best Local Similarity 69.3%; Pred. No. 14;

Matches 9; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 KLVPFATLHERL 13

DB 78 RLVPFATLHERL 90

RESULT 7

ID DADA\_PHOLL STANDARD; PRT; 436 AA.

AC Q7N3Z6;

DT 29-MAR-2004 (Rel. 43, Created)

DT 29-MAR-2004 (Rel. 43, Last sequence update)

DT 25-OCT-2004 (Rel. 45, Last annotation update)

DE D-amino acid dehydrogenase small subunit (EC 1.4.99.1).

GN Name=dadA; OrderedLocNames=plu2561;

OS Photobacterium luminescens (subsp. laumondii).

OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;

OC Enterobacteriaceae; Photobacterium.

OX NCBI\_TaxID=141679;

OX [1]

RP SEQUENCE FROM N.A.

RC STRAIN=TT01;

RX MEDLINE=22957627; PubMed=14528314; DOI=10.1038/nbt886;

RA Duchaud E., Rusniok C., Frangeul L., Buchrieser C., Givaudan A., Taouit S., Bocs S., Boursaux-Eude C., Chandler M., Charles J.-F., Dassa E., Derose R., Derzelle S., Freysinet G., Gaudriault S., Medigue C., Lanois A., Powell K., Signier P., Vincent R., Wingate V., Zouine M., Glaser P., Boemare N., Danchin A., Kunst F.;

RT "The genome sequence of the entomopathogenic bacterium Photobacterium luminescens";

RL Biotechnol. 21:1307-1313(2003).

CC -1- FUNCTION: Oxidative deamination of D-amino acids (By similarity).

CC -1- CATALYTIC ACTIVITY: A D-amino acid + H(2)O + acceptor = a 2-oxo acid + NH(3) + reduced acceptor.

CC -1- COFACTOR: FAD (By similarity).

CC -1- PATHWAY: Alanine catabolism; second (last) step.

CC -1- SUBUNIT: Heterodimer of a small and a large subunit (By similarity).

CC -1- SIMILARITY: Belongs to the dadA oxidoreductase family.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

CC EMBL; BX571867; CAE14935.1; ALT\_INIT.

DR Photobacterium; plu2561; --

DR HAMAP; MF\_01202; --; 1.

DR InterPro; IPR006076; Fad\_oxred.

DR InterPro; IPR000205; NAD\_BS.

DR InterPro; IPR006036; TrkA\_Kuptake.

```

DR Pfam; PF01366; DAO; 1.
DR PRINTS; PR00335; KUP1AKETRKA.
KW Complete proteome; FAD; Flavoprotein; Oxidoreductase.
SQ NP_BIND 3 17 FAD (ADP part) (Potential).
FT SEQUENCE 436 AA; 47971 MW; 24216B33BF0B331C CRC64;

Query Match 63.6%; Score 42; DB 1; Length 436;
Best Local Similarity 61.5%; Pred. No. 59;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 KLVPPFATLHERL 13
DB 422 KLVPPFASHQLHTEL 434

RESULT 8
Q9VLA0 PRELIMINARY; PRT; 4345 AA.
AC Q9VLA0;
ID Q9VLA0;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE CG15828-PA.
GN ORFNames=CG15828;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
[1]
SEQUENCE FROM N.A.
RP MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
RX Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li F.W., Hoskins R.A., Calle R.F.,
RA George R.A., Lewis S.B., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.H., Blazek R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,
RA Abril J.F., Agbayani A., An H.J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Folsler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,
RA Shue B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirkas R., Tector C., Turner E., Venter E., Wang A.H., Wang X.,
RA Wang Z.Y., Wasserman D.A., Weinstock G.M., Weissenbach J.,
RA Williams S.M., Woodage, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,
RA Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195(2000).
[2]
SEQUENCE FROM N.A.
RP MEDLINE=22426065; PubMed=12537568;
RX Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li F.W., Hoskins R.A., Calle R.F.,
RA George R.A., Lewis S.B., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.H., Blazek R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,
RA Abril J.F., Agbayani A., An H.J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Folsler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,
RA Shue B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirkas R., Tector C., Turner E., Venter E., Wang A.H., Wang X.,
RA Wang Z.Y., Wasserman D.A., Weinstock G.M., Weissenbach J.,
RA Williams S.M., Woodage, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,
RA Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195(2000).
[2]
SEQUENCE FROM N.A.
RP MEDLINE=22426065; PubMed=12537568;
RX Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li F.W., Hoskins R.A., Calle R.F.,
RA George R.A., Lewis S.B., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.H., Blazek R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,
RA Abril J.F., Agbayani A., An H.J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Folsler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,
RA Shue B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirkas R., Tector C., Turner E., Venter E., Wang A.H., Wang X.,
RA Wang Z.Y., Wasserman D.A., Weinstock G.M., Weissenbach J.,
RA Williams S.M., Woodage, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,
RA Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195(2000).
[2]
SEQUENCE FROM N.A.
RP MEDLINE=22426070; PubMed=12537573;
RX Kaminker J.S., Bergman C.M., Krommiller B., Carlson J., Svirkas R.,
RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,
RA Ashburner M., Celniker S.E.;
RT "The transposable elements of the Drosophila melanogaster euchromatin:
a genomics perspective."
RL Genome Biol. 3:RESEARCH0084-RESEARCH0084 (2002).
[4]
SEQUENCE FROM N.A.
RP MEDLINE=22426069; PubMed=12537572;
RX Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,
RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
RA Brottier P., Burtis K.C., Celniker S.E., de Grey A.D., Drysdale R.A.,
RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA Lewis S.E.;
RT "Annotation of the Drosophila melanogaster euchromatic genome: a
systematic review."
RL Genome Biol. 3:RESEARCH0083-RESEARCH0083 (2002).
[5]
SEQUENCE FROM N.A.
RP SEQUENCE FROM N.A.
RG FlyBase;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
RN [6]
SEQUENCE FROM N.A.
RG FlyBase;
RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AS003625; AAF52796.1; -.
DR FlyBase; Fggn0032136; CG15828.
DR GO; GO:0005319; Filipid transporter activity; IEA.
DR GO; GO:0006869; Lipid transport; IEA.
DR InterPro; IPR009454; DUF1081.
DR InterPro; IPR001747; Lipid transport N.
DR InterPro; IPR011030; LV superhelical.
DR InterPro; IPR001846; VWF D.
DR Pfam; PF06448; DUF1081; 1.
DR Pfam; PF01347; Vitellogenin_N; 1.
DR Pfam; PF00094; VMD; 1.
DR SMART; SM00638; LPD_N; 1.
DR SMART; SM00216; VMD; 1.
DR SEQUENCE 4345 AA; 495238 MW; 30AAC8439A52165F CRC64;

Query Match 63.6%; Score 42; DB 2; Length 4345;
Best Local Similarity 69.2%; Pred. No. 6e+02;
Matches 9; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 KLVPPFATLHERL 13
DB 3596 KLVPPFGAELNEVL 3608

RESULT 9
Q7KTG2 PRELIMINARY; PRT; 4374 AA.
ID Q7KTG2;
AC Q7KTG2;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE CG15828-PB.
GN ORFNames=CG15828;
OS Drosophila melanogaster (Fruit fly).

```

OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 OX NCBI\_TaxID=7227;

[1]  
 RN SEQUENCE FROM N.A.  
 RP MEDLINE=20196006; PubMed=107311132; DOI=10.1126/science.287.5461.2185;  
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.H., Blazej R.G., Champe M., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,  
 RA Abril J.F., Agbayani A., An H.J., Andrews-Pfannkoch C., Baldwin D.,  
 RA Ballou R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
 RA Borkova D., Botchan M.R., Bouck J., Brockstein P., Brotter P.,  
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrieria S., Fleischmann W.,  
 RA Fouts D.C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,  
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacieb J.M.,  
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Furl V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.Y., Kassaman D.A., Weinstock G.M., Weissbach J.,  
 RA Williams S.M., Woodgett, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,  
 RA Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of Drosophila melanogaster."  
 RL Science 287:2185-2195(2000).  
 [2]  
 RN SEQUENCE FROM N.A.  
 RP MEDLINE=22426065; PubMed=12537568;  
 RA Celniker S.E., Wheeler D.A., Kronmiller B., Carlson J.W., Halpern A.,  
 RA Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,  
 RA George R.A., Hoskins R.A., Lavery T., Muzny D.M., Nelson C.R.,  
 RA Pacleb J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,  
 RA Svirskas R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,  
 RA Weinstock G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.;  
 RT "Finishing a whole-genome shotgun. Release 3 of the Drosophila  
 melanogaster euchromatic genome sequence."  
 RL Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).  
 [3]  
 RN SEQUENCE FROM N.A.  
 RP MEDLINE=22426070; PubMed=12537573;  
 RA Kaminker J.S., Bergman C.M., Kronmiller B., Carlson J., Svirskas R.,  
 RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,  
 RA Ashburner M., Celniker S.E.;  
 RT "The transposable elements of the Drosophila melanogaster euchromatin:  
 a genomics perspective."  
 RL Genome Biol. 3:RESEARCH0084-RESEARCH0084(2002).  
 [4]  
 RN SEQUENCE FROM N.A.  
 RP MEDLINE=22426069; PubMed=12537572;  
 RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,  
 RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,  
 RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,  
 RA Bettencourt B.R., Celniker S.E., de Grey A.D., Drysdale R.A.,

RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,  
 RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,  
 RA Lewis S.E.;  
 RT "Annotation of the Drosophila melanogaster euchromatic genome: a  
 systematic review";  
 RL Genome Biol. 3:RESEARCH0083-RESEARCH0083(2002).  
 [5]  
 RN SEQUENCE FROM N.A.  
 RP FlyBase;  
 RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.  
 [6]  
 RN SEQUENCE FROM N.A.  
 RP FlyBase;  
 RL EMBL; AS003625; AAS64667.1; -;  
 DR GO; GO:0005319; F.lipid transporter activity; IEA.  
 DR GO; GO:0006869; P.lipid transport; IEA.  
 DR InterPro; IPR009454; DUF1081.  
 DR InterPro; IPR001747; Lipid transport N.  
 DR InterPro; IPR011030; LV superhelical.  
 DR InterPro; IPR001846; VWF D.  
 DR Pfam; PF06448; DUF1081; 1.  
 DR Pfam; PF01347; Vitellogenin\_N; 1.  
 DR Pfam; PF00094; VWD; 1.  
 DR SMART; SM00638; LPD\_N; 1.  
 DR SMART; SM00216; VWD; 1.  
 SQ SEQUENCE 4374 AA; 498710 MW; C13F0DB2453D0F95 CRC64;  
 Query Match 53.6%; Score 42; DB 2; Length 4374;  
 Best Local Similarity 69.2%; Pred. NO. 6e+02; Indels 0; Gaps 0;  
 Matches 9; Conservative 1; Mismatches 3;  
 QY 1 KLVPPFATLHERL 13  
 |||||  
 Db 3625 KLVPFQALNEVL 3637  
 RESULT 10  
 O08877 PRELIMINARY; PRT; 258 AA.  
 ID O08877  
 AC O08877; 01-JUL-1997 (TReMBLrel. 04, Created)  
 DT 01-JUL-1997 (TReMBLrel. 04, Last sequence update)  
 DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)  
 DE Apolipoprotein A-I.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=SHR; TISSUE=Spleen;  
 RX MEDLINE=98077648; PubMed=9415807;  
 RA Chiang A.N., Fan K.C., Shaw G.C., Yang U.C.;  
 RT "Repetitive elements in the third intron of murine apolipoprotein A-I  
 gene";  
 RL Biochem. Mol. Biol. Int. 43:989-996(1997).  
 [2]  
 RN SEQUENCE FROM N.A.  
 RC STRAIN=SHR; TISSUE=Spleen;  
 RA Chiang A.-N., Fan K.-C., Shaw G.-C., Yang U.-C.;  
 RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; U75576; AAB58428.1; -;  
 DR HSP; P02647; IAV1.  
 DR GO; GO:0005576; C:extracellular; IEA.  
 DR GO; GO:0008289; F:lipid binding; IEA.  
 DR GO; GO:0006869; P:lipid transport; IEA.  
 DR GO; GO:0042157; P:lipid transport metabolism; IEA.  
 DR InterPro; IPR000074; Apolipoprotein.  
 DR InterPro; IPR009074; Apolipo\_A\_E\_C3.  
 DR Pfam; PF01442; Apolipoprotein; 1.  
 KW Lipoprotein.  
 SQ SEQUENCE 258 AA; 29831 MW; 093FB582E629CDC8 CRC64;





AC Q6P7H6;  
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)  
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)  
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)  
 DE MGC68602 protein.  
 GN Name=MGC68602;  
 OS Xenopus laevis (African clawed frog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;  
 OC Xenopodinae; Xenopus.  
 OX NCBI\_TaxID=8355;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TTSUE=Kidney;  
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,  
 RA Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Pahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
 RA Krzywinski M.I., Skalek U., Smailus D.E., Schnerch A., Schein J.E.,  
 RA Jones S.J., Marra M.A.  
 RA "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences."  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TTSUE=Kidney;  
 RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/dvdy.10174;  
 RA Klein S.B., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,  
 RA Richardson P.;  
 RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus  
 RT initiative."  
 RL Dev. Dyn. 225:384-391(2002).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC TTSUE=Kidney;  
 RA Klein S., Strausberg R.;  
 RL Submitted (NOV-2003) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: BC061667; AAH61667.1;  
 DR GO: GO:0005576; C:extracellular; IEA.  
 DR GO: GO:0008289; F:lipid binding; IEA.  
 DR GO: GO:0006869; F:lipid transport; IEA.  
 DR GO: GO:0042157; P:lipoprotein metabolism; IEA.  
 DR InterPro: IPR000074; Apolipoprotein.  
 DR InterPro: IPR009074; Apolipo\_A\_E\_C3.  
 DR Pfam: PF01442; Apolipoprotein; 2.  
 SQ SEQUENCE 367 AA; 42332 MW; 699D578FE4626E78 CRC64;  
 Query Match 62.1%; Score 41; DB 2; Length 367;  
 Best Local Similarity 38.5%; Pred. No. 76;  
 Matches 5; Conservative 7; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 KLVPFATELHERL 13  
 DB 79 QLIPFAKQIHQDM 91  
 RESULT 15  
 Q89EW7 PRELIMINARY; PRT; 897 AA.  
 AC Q89EW7;  
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)

DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
 DE Molybdenum ABC transporter ATP-binding protein.  
 GN Name=modC; OrderedLocusNames=blr6953;  
 OS Bradyrhizobium japonicum.  
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;  
 OC Bradyrhizobiaceae; Bradyrhizobium.  
 OX NCBI\_TaxID=375;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=USD110;  
 RX MEDLINE=22484998; PubMed=12597275;  
 RA Kaneko T., Nakamura Y., Sato S., Minamisawa K., Uchiumi T.,  
 RA Sasamoto S., Watanabe A., Idesawa K., Iriguchi M., Kawashima K.,  
 RA Kohara M., Matsumoto M., Shimo S., Tsuruoka H., Wada T., Yamada M.,  
 RA Tabata S.;  
 RT "Complete genomic sequence of nitrogen-fixing symbiotic bacterium  
 RT Bradyrhizobium japonicum USDA110."  
 RL DNA Res. 9:189-197(2002).  
 CC -1- SIMILARITY: Belongs to the ABC transporter family.  
 DR EMBL: AP005960; BAC52218.1;  
 DR HSSP: Q9YGA6; 1G29.  
 DR GO: GO:0016020; C:membrane; IEA.  
 DR GO: GO:0005524; F:ATP binding; IEA.  
 DR GO: GO:0042626; F:ATPase activity, coupled to transmembrane m.; IEA.  
 DR GO: GO:0000166; F:nucleotide binding; IEA.  
 DR GO: GO:0006810; P:transport; IEA.  
 DR InterPro: IPR003593; AAA\_ATPase.  
 DR InterPro: IPR003439; ABC\_transporter.  
 DR InterPro: IPR008995; MOP\_like.  
 DR InterPro: IPR005116; TOBE.  
 DR Pfam: PF00005; ABC\_tran; 1.  
 DR Pfam: PF03459; TOBE; 1.  
 DR ProDom: PD000006; ABC\_transporter; 1.  
 DR SMART: SM00382; AAA; 1.  
 DR PROSITE: PS00211; ABC\_TRANSPORTER\_1; 1.  
 DR PROSITE: PS00893; ABC\_TRANSPORTER\_2; 1.  
 KW ATP-binding; Complete proteome.  
 SQ SEQUENCE 897 AA; 96173 MW; 154C54F8D41A4D12 CRC64;  
 Query Match 62.1%; Score 41; DB 2; Length 897;  
 Best Local Similarity 53.8%; Pred. No. 1.9e+02;  
 Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;  
 QY 1 KLVPFATELHERL 13  
 DB 173 EILPFLERLHERL 185  
 RESULT 16  
 Q7Q9K3 PRELIMINARY; PRT; 1056 AA.  
 ID Q7Q9K3  
 AC Q7Q9K3; (TrEMBLrel. 26, Created)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
 DE AGCP15000 (fragment).  
 GN Name=agCG46168; ORFNames=ENSANGG00000007942;  
 OS Anopheles gambiae str. PEST.  
 OC Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Diptera; Nematocera; Anopheles.  
 OX NCBI\_TaxID=180454;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=PEST;  
 RA Anopheles Genome Sequencing Consortium;  
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.  
 CC -1- CAUTION: The sequence shown here is derived from an  
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is  
 CC preliminary data.  
 DR EMBL: AAA01008900; EAA09356.1; -.  
 FT NON\_TER 1  
 FT NON\_TER 1056 1056

```

SQ SEQUENCE 1056 AA; 121746 MW; 191129F749BA27A2 CRC64;
Query Match 62.1%; Score 41; DB 2; Length 1056;
Best Local Similarity 61.5%; Pred. No. 2.2e+02;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 KLVFPFATELHERL 13
Db 61 KLQPFANKLHQL 73

RESULT 17
Q6NJ57 PRELIMINARY; PRT; 1179 AA.
AC Q6NJ57;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE Putative FtsK/SpoIIIE family ATP-binding protein.
GN OrderedLocusNames=DIP0556;
OS Corynebacterium diphtheriae.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Corynebacteriaceae; Corynebacterium.
OX NCBI_TaxID=1717;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Biotype Gravis / NCTC 13129;
RX MEDLINE=22965443; PubMed=14602910; DOI=10.1093/nar/gkg874;
RA Cerdeno-Tarraga A.-M., Estratiou A., Dover L.G., Holden M.T.G.,
RA Fallen M.J., Bentley S.D., Besra G.S., Churcher C.M., James K.D.,
RA De Zoysa A., Chillingworth T., Cronin A., Dowd L., Feltwell T.,
RA Hamlin N., Hollroyd S., Jagels K., Moule S., Quail M.A.,
RA Rabinowitz E., Rutherford K.M., Thomson N.R., Unwin L.,
RA Whitehead S., Barrell B.G., Parkhill J.;
RT "The complete genome sequence and analysis of Corynebacterium
RT diphtheriae NCTC13129."
RL Nucleic Acids Res. 31:6516-6523 (2003).
DR EMBL; BX249355; CAE49068.1; -.
DR GO; GO:000166; F:nucleotide binding; IEA.
DR InterPro; IPR003593; AAA_ATPase.
DR IPR002543; FtsK_SpoIIIE.
DR Pfam; PF01580; FtsK_SpoIIIE; 2.
DR SMART; SM00382; AAA; 3.
DR PROSITE; PS50901; FTSK; 3.
KW ATP-binding; Complete proteome.
SQ SEQUENCE 1179 AA; 128991 MW; 2939055714E35A69 CRC64;
Query Match 62.1%; Score 41; DB 2; Length 1179;
Best Local Similarity 72.7%; Pred. No. 2.4e+02;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 KLVFPFATELHE 11
Db 931 KLLPTAVELHE 941

RESULT 18
TNF5_RAT STANDARD; PRT; 260 AA.
AC Q922V2; Q9R254;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Tumor necrosis factor ligand superfamily member 5 (CD40 ligand) (CD40-
DE L).
GN Names=Tnf5; Synonyms=Cd40l, Cd40lg;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley; TISSUE=Splenocyte;

RX MEDLINE=20284949; PubMed=10826698;
RA Hallatt K.M., Oaks M.K.;
RT "Nucleotide sequence of the rat CD40 ligand.";
RL DNA Seq. 10:405-406 (2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=PVG; TISSUE=Spleen;
RA Daniel K.C., Foss Y., Mousavi A., Macary P., Kemeny D.M.,
RA Farzaneh F., Gaken J.A.;
RT "Cloning and sequencing of rat CD40 ligand.";
RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Cytokine that binds to TNFRSF5. Mediates B-cell
CC proliferation in the absence of co-stimulus as well as IGE
CC production in the presence of IL-4. Involved in immunoglobulin
CC class switching (By similarity).
CC -!- SUBUNIT: Homotrimer (By similarity).
CC -!- SUBCELLULAR LOCATION: Type II membrane protein. Also exists as an
CC extracellular soluble form (By similarity).
CC -!- PTM: The soluble form derives from the membrane form by
CC proteolytic processing (By similarity).
CC -!- SIMILARITY: Belongs to the tumor necrosis factor family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF116582; AAD22460.1; -.
CC EMBL; AF013985; AAD09323.1; -.
CC HSSP; P29965; ITALY.
CC RGD; 708418; LOC84349.
CC GO; GO:0016021; C:integral to membrane; ISS.
CC GO; GO:0003174; F:CD40 receptor binding; ISS.
CC GO; GO:0042100; P:B-cell proliferation; ISS.
CC GO; GO:0006954; P:inflammatory response; ISS.
CC GO; GO:0007159; P:leukocyte cell adhesion; ISS.
CC GO; GO:0030168; P:platelet activation; ISS.
CC InterPro; IPR003263; TNF_5.
CC InterPro; IPR006052; TNF_family.
CC InterPro; IPR008983; TNF_like.
CC Pfam; PF00229; TNF; 1.
CC InterPro; IPR003636; TNF_subf.
CC PRINTS; PR01702; CD40LIGAND.
CC ProDom; PD008600; TNF_5; 1.
CC ProDom; PD002012; TNF_subf; 1.
CC SMART; SM00207; TNF; 1.
CC PROSITE; PS00251; TNF_1; 1.
CC PROSITE; PS50049; TNF_2; 1.
KW Cytokine; Glycoprotein; Signal-anchor; Transmembrane.
FT CHAIN 1 260 Tumor necrosis factor ligand superfamily
FT member 5, membrane form.
FT Tumor necrosis factor ligand superfamily
FT member 5, soluble form (By similarity).
FT Cytoplasmic (Potential).
FT Signal-anchor for type II membrane
FT protein (Potential).
FT Extracellular (Potential).
FT Cleavage (By similarity).
FT Potential.
FT N-linked (GlcNAc...) (Potential).
FT P -> L (in Ref. 1).
FT R -> K (in Ref. 1).
FT CONFLICT 114 114 I -> F (in Ref. 1).
FT CONFLICT 255 255 I -> F (in Ref. 1).
SQ SEQUENCE 260 AA; 29259 MW; B3D3757DE60DB73A CRC64;
Query Match 60.6%; Score 40; DB 1; Length 260;
Best Local Similarity 72.7%; Pred. No. 81;
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 VFPATELHERL 13

```

||||| ||| |||  
40 VPPAVLHREL 50

Db

## RESULT 19

APAI RABIT  
ID APAI\_RABIT STANDARD; PRT; 266 AA.  
AC P09809;  
DT 01-MAR-1989 (Rel. 10, Created)  
DT 01-JAN-1990 (Rel. 13, Last sequence update)  
DT 05-JUL-2004 (Rel. 44, Last annotation update)  
DE Apolipoprotein A-I precursor (Apo-AI).  
GN Name=APOA1;  
OS Oryctolagus cuniculus (Rabbit).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Legomorphia; Leporidae; Oryctolagus.  
OX NCBI\_TaxID=9986;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=22AP AI; TISSUE=Small intestine;  
RA Paraskevopoulou T.B., Kritis A., Zannis V.I.;  
RL Submitted (JUL-1989) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Intestine;  
RX MEDLINE=88082866; PubMed=3121329;  
RA Pan T.C., Hao Q.L., Yamin T.T., Dai P.H., Chen B.S., Chen S.L.,  
RA Kroon P.A., Chao Y.S.;  
RT "Rabbit apolipoprotein A-I mRNA and gene. Evidence that rabbit  
RT apolipoprotein A-I is synthesized in the intestine but not in the  
RT liver.";  
RL Eur. J. Biochem. 170:99-104 (1987).  
RN [3]  
RP SEQUENCE OF 25-266.  
RX MEDLINE=87030294; PubMed=3095115;  
RA Yang C., Yang T., Pownall H.J., Gotto A.M. Jr.;  
RT "The primary structure of apolipoprotein A-I from rabbit high-density  
RT lipoprotein";  
RL Eur. J. Biochem. 160:427-431 (1986).  
CC -1- FUNCTION: Participates in the reverse transport of cholesterol  
CC from tissues to the liver for excretion by promoting cholesterol  
CC efflux from tissues and by acting as a cofactor for the lecithin  
CC cholesterol acyltransferase (LCAT).  
CC -1- SUBCELLULAR LOCATION: Secreted.  
CC -1- TISSUE SPECIFICITY: Major protein of plasma HDL, also found in  
CC chylomicrons.  
CC -1- SIMILARITY: Belongs to the apolipoprotein A1/A4/E family.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC -----  
CC EMBL; X15908; CAA34024.1; --  
CC EMBL; X06658; CAA29857.1; --  
CC EMBL; X06659; CAA29858.1; --  
CC PIR; S06064; LPRB12.  
CC HSP; P02647; IAV1.  
CC InterPro; IPR009074; Apolipo\_A\_E\_C3.  
CC InterPro; IPR000074; Apolipoprotein.  
CC Pfam; PF01442; Apolipoprotein; 1.  
CC Cholesterol metabolism; Direct protein sequencing; HDL;  
KW Lipid transport; Plasma; Repeat; Signal.  
FT SIGNAL 1 18  
FT PROPEP 19 24  
FT CHAIN 25 266 Apolipoprotein A-I.  
FT DOMAIN 67 266 10 X approximate tandem repeats.  
FT REPEAT 67 88 1.  
FT REPEAT 89 110 2.  
FT REPEAT 111 121 3 (half-length).  
FT REPEAT

FT REPEAT 122 143 4.  
FT REPEAT 144 165 5.  
FT REPEAT 166 187 6.  
FT REPEAT 188 209 7.  
FT REPEAT 210 231 8.  
FT REPEAT 232 242 9 (half-length).  
FT REPEAT 243 266 10.  
FT CONFLICT 18 18 A -> R (in Ref. 2; CAA29858).  
FT CONFLICT 44 44 Missing (in Ref. 3).  
FT CONFLICT 45 45 V -> I (in Ref. 2; CAA29858).  
FT CONFLICT 107 107 E -> Q (in Ref. 3).  
FT CONFLICT 123 123 Y -> F (in Ref. 2; CAA29857).  
FT CONFLICT 147 147 A -> V (in Ref. 2; CAA29858 and 3).  
FT CONFLICT 150 150 R -> G (in Ref. 2; CAA29858/CAA29857).  
FT CONFLICT 191 191 N -> Q (in Ref. 3).  
FT CONFLICT 195 195 Missing (in Ref. 2; CAA29858/CAA29857).  
FT CONFLICT 211 211 S -> K (in Ref. 2; CAA29858/CAA29857).  
FT CONFLICT 255 255 VL -> LV (in Ref. 3).  
FT CONFLICT 256 256 L -> V (in Ref. 2; CAA29858/CAA29857).  
SQ SEQUENCE 266 AA; 30591 MW; 0FF6DB386497C7D2 CRC64;

Query Match 60.6%; Score 40; DB 1; Length 266;

Best Local Similarity 53.8%; Pred No. 83;

Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 KLVPFATELHERL 13

DB 185 KLAPYSNELQRL 197

## RESULT 20

Q66IL6 PRELIMINARY; PRT; 368 AA.

AC Q66IL6; 2004 (TrEMBLrel. 28, Created)

DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)

DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)

DE Hypothetical protein.

OS Xenopus tropicalis (Western clawed frog) (Silurana tropicalis).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;

OC Xenopodinae; Xenopus.

OX NCBI\_TaxID=8364;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Whole body;

RX PubMed=12477932; DOI=10.1073/pnas.242603899;

RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,

RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,

RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,

RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haieh F.,

RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,

RA Scapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E.,

RA Brownstein M.J., Udén T.B., Toehiyuki S., Carninci P., Prange C.,

RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,

RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,

RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,

RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,

RA Faney J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,

RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,

RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,

RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,

RA Krzywinski M.I., Skaleka U., Smallos D.E., Schnerch A., Schein J.E.,

RA Jones S.J., Marra M.A.;

RT "Generation and initial analysis of more than 15,000 full-length human

RT and mouse cDNA sequences.";

RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).

RN [2]

RP SEQUENCE FROM N.A.

RC TISSUE=Whole body;

RA Klein S., Gerhard D.S.;

RL Submitted (AUG-2004) to the EMBL/GenBank/DBJ databases.

EMBL; BC081300; AAH81300.1; --.

```

DR InterPro; IPR000074; Apolipoprotein.
DR InterPro; IPR009074; Apolipo A E C3.
DR Pfam; PF01442; Apolipoprotein; 1.
KW Hypothetical protein.
SQ SEQUENCE 368 AA; 42501 MW; C4EA22D395FC9DB4 CRC64;

  Query Match      60.6%; Score 40; DB 2; Length 368;
  Best Local Similarity 38.5%; Pred. No. 1.2e+02;
  Matches 5; Conservative 7; Mismatches 1; Indels 0; Gaps 0;

QY 1 KLVPPFATELHERL 13
   :::|||::|:|:|
DB 79 QVTFPAKIHQDL 91

RESULT 21
Q6CGG7 PRELIMINARY; PRT; 585 AA.
AC Q6CGG7;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Similar to tr|Q06436 Saccharomyces cerevisiae YJR427w.
GN ORFNames=YAL10A19149;
OS Yarrowia lipolytica CL1B99.
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Dipodascaceae; Yarrowia.
OX NCBI_TaxID=284591;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CL1B99;
RG Genolevures;
RA Lafontaine I., de Montigny J., March C., Neuveglise C., Talla E.,
RA Goffard N., Frangeul L., Aigle M., Anthouard V., Babour A., Barbe V.,
RA Barnay S., Blanchin S., Beckerich J.M., Beyne E., Bleykasten C.,
RA Boissrame A., Boyer J., Cattolico L., Confanioleri F., de Daruvar A.,
RA Despons L., Fabre E., Fairhead C., Ferry-Dumazet H., Groppi A.,
RA Hantraye F., Hennequin C., Jaumiaux N., Joyet P., Kachouri R.,
RA Kerrest A., Koszul R., Lemaire M., Lesur I., Ma L., Muller H.,
RA Nicaud J.M., Nikolski M., Oztas S., Ozier-Kalogeropoulos O.,
RA Pellenn S., Potier S., Richard G.F., Straub M.L., Suleau A.,
RA Pellenn D., Tekala F., Wesolowski-Louvel M., Westhof E., Wirth B.,
RA Zeniou-Meyer M., Zivanovic I., Bolotin-Fukuhara M., Thierry A.,
RA Bouchier C., Caudron B., Scarpelli C., Gaillardin C., Weissenbach J.,
RA Wincker P., Souciet J.L.;
RT "Genome evolution in yeasts.";
RL Nature 430:35-44(2004).
[2]
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CL1B99;
RG Genoscope;
RA Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Contains 1 RING-type zinc finger.
DR EMBL; CR382127; CAG84183.1; -.
DR GO; GO:0000151; C:ubiquitin ligase complex; IEA.
DR GO; GO:0004842; F:ubiquitin-protein ligase activity; IEA.
DR GO; GO:0008270; F:zinc ion binding; IEA.
DR GO; GO:0016567; P:protein ubiquitination; IEA.
DR InterPro; IPR001841; Znf_ring.
DR Pfam; PF00097; zf-C3HC4; 1.
DR SMART; SM00184; RING; 1.
DR PROSITE; PS00518; ZF_RING 1; 1.
DR PROSITE; PS50089; ZF_RING 2; 1.
KW Metal-binding; Zinc; Zinc-finger.
SQ SEQUENCE 585 AA; 65101 MW; BAA6D587D0F92CD2 CRC64;

  Query Match      60.6%; Score 40; DB 2; Length 585;
  Best Local Similarity 66.7%; Pred. No. 1.8e+02;
  Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 KLVPPFATELHER 12
   || ||| ||| :|
DB 79 QVTFPAKIHQDL 91

Db 439 KLVPPFATELHER 450
RESULT 22
Q7ULM6 PRELIMINARY; PRT; 656 AA.
ID Q7ULM6;
AC Q7ULM6;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=RB9405;
OS Rhodopirellula baltica.
OC Bacteria; Planctomycetes; Planctomycetacia; Planctomycetales;
OC Planctomycetaceae; Pirellula.
OX NCBI_TaxID=117;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1;
RX MEDLINE=22735913; PubMed=12835416; DOI=10.1073/pnas.1431443100;
RA Gloeckner F.O., Kube M., Bauer M., Teeling H., Lombardot T.,
RA Ludwig W., Gade D., Beck A., Borzym K., Heitmann K., Rabus R.,
RA Schlesner H., Amann R., Reinhardt R.;
RA "Complete genome sequence of the marine planctomycete Pirellula sp.
RT strain 1.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:8298-8303(2003).
DR EMBL; BX294149; CAD76243.1; -.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 656 AA; 75050 MW; 977C8888D2510F02 CRC64;

  Query Match      60.6%; Score 40; DB 2; Length 656;
  Best Local Similarity 70.0%; Pred. No. 2.1e+02;
  Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 VVPFATELHER 12
   :||| ||| |||
DB 255 LPFTSLHER 264

RESULT 23
Q7N819 PRELIMINARY; PRT; 150 AA.
ID Q7N819;
AC Q7N819;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Similar to unknown protein.
DE OrderedLocusNames=plu0929;
GN Photorhabdus luminescens (subsp. laumondii).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Photorhabdus.
OX NCBI_TaxID=141679;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TT01;
RX MEDLINE=22957627; PubMed=14528314;
RA Duchaud E., Rusniok C., Frangeul L., Buchrieser C., Givaudan A.,
RA Taourit S., Bocs S., Boursaux-Eude C., Chandler M., Charles J.-F.,
RA Dassea E., Derose R., Derzelle S., Freyssinet G., Gaudriault S.,
RA Medigue C., Lanois A., Powell K., Signier P., Vincent R., Wingate V.,
RA Zouline M., Glaser P., Boemare N., Danchin A., Kunst F.;
RT "The genome sequence of the entomopathogenic bacterium Photorhabdus
RT luminescens.";
RL Nat. Biotechnol. 21:1307-1313(2003).
DR EMBL; BX571862; CAE13224.1; -.
DR PhotoList; plu0929; -.
DR InterPro; IPR002716; PiliT_N.
DR Pfam; PF01850; PIN; 1.
KW Complete proteome.
SQ SEQUENCE 150 AA; 17131 MW; 7FBE5C3A2B05F3D8 CRC64;

  Query Match      59.1%; Score 39; DB 2; Length 150;
  Best Local Similarity 50.0%; Pred. No. 71;

```

Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;  
QY 1 KLVPPFATELHER 12  
DB 86 KIVPLTVEVHDR 97

## RESULT 24

Q8GF72 PRELIMINARY; PRT; 177 AA.  
ID Q8GF72  
AC Q8GF72;  
DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
DE Orf75.  
GN Name=orf75;  
OS Photorhabdus luminescens (Xenorhabdus luminescens).  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
OC Enterobacteriaceae; Photorhabdus.  
OX NCBI\_TaxID=29488;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=W14;  
RX MEDLINE=21185117; PubMed=11286884; DOI=10.1016/S0966-842X(01)01978-3;  
RA Waterfield N.R., Bowen D.J., Fetherston J.D., Perry R.D.,  
RA french-Constant R.H.;  
RT "The tc genes of Photorhabdus: a growing family";  
RL Trends Microbiol. 9:185-191(2001).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=W14;  
RX MEDLINE=22454990; PubMed=12564983; DOI=10.1016/S0966-842X(02)02463-0;  
RA Waterfield N.R., Daborn P.J., french-Constant R.H.;  
RT "Genomic islands in Photorhabdus";  
RL Trends Microbiol. 10:541-545(2002).  
DR EMBL; AF346500; AAO17229.1; -.  
DR InterPro; IPR002716; Pilt N.  
DR InterPro; IPR006596; Pinc.  
DR Pfam; PF01850; PIN; 1.  
DR SMART; SM00670; PIN; 1.  
SQ SEQUENCE 177 AA; 19857 MW; 31C279D8D49172B9 CRC64;

Query Match 59.1%; Score 39; DB 2; Length 177;  
Best Local Similarity 50.0%; Pred. No. 84;  
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 KLVPPFATELHER 12  
DB 113 KIVPLTVEVHDR 124

## RESULT 25

Q830L3 PRELIMINARY; PRT; 220 AA.  
ID Q830L3  
AC Q830L3;  
DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
DE Transcriptional regulator.  
GN OrderedLocusNames=EF2767;  
OS Enterococcus faecalis (Streptococcus faecalis).  
OC Bacteria; Firmicutes; Lactobacillales; Enterococcaceae; Enterococcus.  
OX NCBI\_TaxID=1351;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=V583 / ATCC 700802;  
RX MEDLINE=22550857; PubMed=12663927; DOI=10.1126/science.1080613;  
RA Paulsen I.T., Banerjee L., Myers G.S.A., Nelson K.E., Seshadri R.,  
RA Read T.D., Fouts D.E., Eisen J.A., Gill S.R., Heidelberg J.F.,  
RA Tettelin H., Dodson R.J., Umayam L.A., Brinkac L.M., Beanan M.J.,  
RA Daugherty S.C., DeBoy R.T., Durkin S.A., Kolonay J.F., Madupu R.,  
RA Nelson W.C., Vamathevan J.J., Tran B., Upton J., Hansen T., Shetty J.,  
RA Khouri H.M., Utterback T.R., Radune D., Ketchum K.A., Dougherty B.A.,

RA Fraser C.M.;  
RT "Role of mobile DNA in the evolution of vancomycin-resistant  
RL Enterococcus faecalis";  
RL Science 299:2071-2074(2003).  
DR EMBL; AE016955; AAO82465.1; -.  
DR TIGR; EF2767; -.  
DR InterPro; IPR004305; TENA\_THI-4.  
DR Pfam; PF03070; TENA\_THI-4; 1.  
SQ Complete proteome.  
KW SEQUENCE 220 AA; 25626 MW; 1024606E58E24D4F CRC64;

Query Match 59.1%; Score 39; DB 2; Length 220;  
Best Local Similarity 87.5%; Pred. No. 1e+02;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 PFATELHE 11  
DB 22 PFITELHE 29

## RESULT 26

P87210 PRELIMINARY; PRT; 353 AA.  
ID P87210  
AC P87210;  
DT 01-JUL-1997 (TrEMBLrel. 04, Created)  
DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
DE MEI3.  
GN Name=mei-3;  
OS Neurospora crassa.  
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;  
OC Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.  
OX NCBI\_TaxID=5141;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=74-OR23-1A; PubMed=8552049;  
RX MEDLINE=96133684; PubMed=8552049;  
RA Hatakeyama S., Ishii C., Inoue H.;  
RT "Identification and expression of the Neurospora crassa mei-3 gene  
RT which encodes a protein homologue to Rad51 of Saccharomyces  
cerevisiae";  
RL Mol. Gen. Genet. 249:439-446(1995).  
CC -1- SIMILARITY: Belongs to the reca family.  
DR EMBL; D29638; BAA20366.1; -.  
DR PIR; S70629; S70629.  
DR HSP; Q06609; INOW.  
DR GO; GO:0005622; C:intracellular; IEA.  
DR GO; GO:0005524; F:ATP binding; IEA.  
DR GO; GO:0003684; F:damaged DNA binding; IEA.  
DR GO; GO:0008094; F:DNA-dependent ATPase activity; IEA.  
DR GO; GO:0000166; F:nucleotide binding; IEA.  
DR GO; GO:0006310; P:DNA recombination; IEA.  
DR GO; GO:0006281; P:DNA repair; IEA.  
DR InterPro; IPR003593; AAA\_ATPase.  
DR InterPro; IPR000445; HhH.  
DR InterPro; IPR003583; HhH 1.  
DR InterPro; IPR010995; Rad51\_N.  
DR InterPro; IPR001553; RecA.  
DR Pfam; PF00633; HhH; 1.  
DR ProDom; PD000229; RecA; 1.  
DR SMART; SM00382; AAA; 1.  
DR SMART; SM00278; HhH1; 1.  
DR PROSITE; PS0162; RECA\_2; 1.  
DR PROSITE; PS0163; RECA\_3; 1.  
KW ATP-binding.  
SQ SEQUENCE 353 AA; 38202 MW; B79DC4D2DA0B8548 CRC64;

Query Match 59.1%; Score 39; DB 2; Length 353;  
Best Local Similarity 56.2%; Pred. No. 1.7e+02;  
Matches 9; Conservative 2; Mismatches 1; Indels 4; Gaps 1;

QY 1 KLVPPF---ATELHER 12  
|||||  
|||||

```

Db      84 KLVPMTGTTATEMHQR 99

RESULT 27
Q9CA19
ID      Q9CA19      PRELIMINARY;      PRT;      697 AA.
AC      Q9CA19;
DT      01-JUN-2001 (TrEMBLrel. 17, Created)
DT      01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT      01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE      Hypothetical protein F28922.5.
GN      Name=F28922.5;
OS      Arabidopsis thaliana (Mouse-ear cress).
OC      Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC      Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC      eurosoids II; Brassicales; Brassicaceae; Arabidopsis.
OX      NCBI_TaxID=3702;
RN      [1]
RP      SEQUENCE FROM N.A.
RA      Lin X., Kaul S., Town C.D., Benito M.-I., Creasy T.H., Haas B.J.,
RA      Wu D., Maiti R., Ronning C.M., Koo H., Fujii C.Y., Utterback T.R.,
RA      Barnstead M.E., Bowman C.L., White O., Nierman W.C., Fraser C.M.,
RA      Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
RL      [2]
RN      SEQUENCE FROM N.A.
RA      Town C.D., Kaul S.,
RA      Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
RL      -!- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
CC      EMBL; AC010926; AGS1852.1; -.
DR      PIR; E96752; E96752.
DR      HSP; Q62838; 1LUF.
DR      GO; GO:0005524; F:ATP binding; IEA.
DR      GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR      GO; GO:0016740; F:transferase activity; IEA.
DR      GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR      InterPro; IPR011009; Kinase like.
DR      InterPro; IPR000719; Prot kinase.
DR      InterPro; IPR008271; Ser Thr pkin_AS.
DR      ProDom; PD000001; Prot kinase; 1.
DR      PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR      PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR      ATP-binding; Hypothetical protein; Kinase.
KW      Serine/threonine-protein kinase; Transferase.
SQ      SEQUENCE 697 AA; 77917 MW; 67C01D858364BE5 CRC64;

Query Match      59.1%; Score 39; DB 2; Length 697;
Best Local Similarity 61.5%; Pred. No. 3.3e+02;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY      1 KLVPMTGTTATEMHQR 13
Db      647 RLKEFATEDHERI 659

RESULT 28
Q7R3A9
ID      Q7R3A9      PRELIMINARY;      PRT;      2710 AA.
AC      Q7R3A9;
DT      01-MAR-2004 (TrEMBLrel. 26, Created)
DT      01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT      01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE      GLP 111 35594 43726.
GN      Giardia lamblia ATCC 50803.
OS      Eukaryota; Diplomonadida; Hexamitidae; Giardia.
OX      NCBI_TaxID=184922;
RN      [1]
RP      SEQUENCE FROM N.A.
RA      STRAIN=WB C6;
RA      Morrison H.G., McArthur A.G., Adam R.D., Aley S.B., Gillin F.D.,
RA      Olsen G.J., Sogin M.L.;
RA      "draft sequence of the Giardia lamblia genome.";
RL      Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
RL      -!- CAUTION: The sequence shown here is derived from an

Cc      EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
Cc      preliminary data.
DR      EMBL; AACB0100016; EAA41809.1; -.
DR      GO; GO:0030286; C:dynenin complex; IEA.
DR      GO; GO:0005622; C:intracellular; IEA.
DR      GO; GO:0005524; F:ATP binding; IEA.
DR      GO; GO:0003777; F:microtubule motor activity; IEA.
DR      GO; GO:0007018; P:microtubule-based movement; IEA.
DR      GO; GO:0006810; P:transport; IEA.
DR      InterPro; IPR004273; Dynein_heavy.
DR      InterPro; IPR000897; SRP54.
DR      Pfam; PF03028; Dynein_heavy; 1.
DR      PROSITE; PS00300; SRP54; 35CC068501F67345 CRC64;
SQ      SEQUENCE 2710 AA; 303764 MW; 35CC068501F67345 CRC64;

Query Match      59.1%; Score 39; DB 2; Length 2710;
Best Local Similarity 54.5%; Pred. No. 1.3e+03;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY      1 KLVPMTGTTATEMHQR 11
Db      1022 RLAPATEIHD 1032

RESULT 29
Q48926
ID      Q48926      PRELIMINARY;      PRT;      2796 AA.
AC      Q48926;
DT      01-NOV-1996 (TrEMBLrel. 01, Created)
DT      01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT      01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE      Fatty acid synthase.
GN      Name=fas;
OS      Mycobacterium bovis.
OC      Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC      Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX      NCBI_TaxID=1765;
RN      [1]
RP      SEQUENCE FROM N.A.
RA      STRAIN=BCG-Pasteur;
RA      MEDLINE=96200863; PubMed=8621098; DOI=10.1016/0378-1119(95)00842-X;
RA      Fernandes N.D., Kolattukudy P.E.;
RT      "Cloning, sequencing and characterization of a fatty acid synthase-
RT      encoding gene from Mycobacterium tuberculosis var. bovis BCG.";
RL      Gene 170:95-99(1996).
DR      EMBL; U36763; AAB03809.1; -.
DR      PIR; JC4743; JC4743.
DR      GO; GO:0005835; C:fatty-acid synthase complex; IEA.
DR      GO; GO:0004312; F:fatty-acid synthase activity; IEA.
DR      GO; GO:0016491; F:oxidoreductase activity; IEA.
DR      GO; GO:0016740; F:transferase activity; IEA.
DR      GO; GO:0006633; P:fatty acid biosynthesis; IEA.
DR      GO; GO:0008152; P:metabolism; IEA.
DR      InterPro; IPR001227; Ac_transferase.
DR      InterPro; IPR003965; Fatty_acid_synth.
DR      InterPro; IPR000794; Ketoacyl_synth.
DR      InterPro; IPR002539; MaoC_dehydratas.
DR      Pfam; PF00698; Acyl_transf_1; 1.
DR      Pfam; PF00109; ketoacyl-synt; 1.
DR      Pfam; PF02801; ketoacyl-synt_C; 1.
DR      Pfam; PF01575; MaoC_dehydratas; 1.
DR      PRINTS; PR01483; FASYNTHASE.
DR      PROSITE; PS00606; B_KETOACYL_SYNTHASE; UNKNOWN_1.
KW      Transferase.
SQ      SEQUENCE 2796 AA; 303676 MW; 47B87169126D2504 CRC64;

Query Match      59.1%; Score 39; DB 2; Length 2796;
Best Local Similarity 80.0%; Pred. No. 1.3e+03;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      4 PFATELHERL 13
Db      2755 PFATELHERL 2764

```

```

RESULT 30
Q7D623 PRELIMINARY; PRT; 3069 AA.
ID Q7D623
AC Q7D623
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Fatty-acid synthase.
GN OrderedLocusNames=MT2600;
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H37RV;
RX MEDLINE=98295987; PubMed=9634230; DOI=10.1038/31159;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C.M.,
RA Harris D.E., Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III,
RA Tekala F., Badcock K., Basham D., Brown D., Chillingworth T.,
RA Connor R., Davies R.M., Devlin K., Feltwell T., Gentles S., Hamlin N.,
RA Holroyd S., Hornsby T., Jagels K., Krogh A., McLean J., Moule S.,
RA Murphy L.D., Oliver S., Osborne J., Quail M.A., Rajandream M.A.,
RA Rogers J., Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence.";
RL Nature 393:537-544(1998).
DR EMBL; BX842580; CAB06201.1; -.
DR PIR; H70656; H70656.
DR TubercuList; RV2524c; -.
DR GO; GO:0005835; C:fatty-acid synthase complex; IEA.
DR GO; GO:0004312; F:fatty-acid synthase activity; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0006633; P:fatty acid biosynthesis; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR001227; Ac_transferase.
DR InterPro; IPR003965; Fatty_acid_synth.
DR InterPro; IPR000794; Ketoacyl_synth.
DR InterPro; IPR002539; Maoc_dehydratas.
DR Pfam; PF00698; Acyl_transf_1; 1.
DR Pfam; PF00109; ketoacyl-synt; 1.
DR Pfam; PF02801; Ketoacyl-synt_C; 1.
DR Pfam; PF01575; Maoc_dehydratas; 1.
DR PRINTS; PR01483; FASYNTHASE.
DR PROSITE; PS00606; B_KETOACYL_SYNTHASE; UNKNOWN 1.
KW Acyltransferase; Complete proteome; Transferase.
SQ SEQUENCE 3069 AA; 326251 MW; 25EF022D54943D4C CRC64;

Query Match 59.1%; Score 39; DB 2; Length 3069;
Best Local Similarity 80.0%; Pred. No. 1.5e+03;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 PFATLHERL 13
Db 2881 PNETLHERL 2890

RESULT 32
Q7TYD8 PRELIMINARY; PRT; 3069 AA.
ID Q7TYD8
AC Q7TYD8
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE PROBABLE FATTY ACID SYNTHASE FAS (FATTY ACID SYNTHETASE) (EC 2.3.1.-
DE ).
GN Name=fas; OrderedLocusNames=MD2553c;
OS Mycobacterium bovis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1765;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AF2122/97;
RX MEDLINE=22709107; PubMed=12788972; DOI=10.1073/pnas.1130426100;
RA Garner T., Eiglmeier K., Camus J.-C., Medina N., Mansoor H.,
RA Pryor M., Duthoy S., Grondin S., Lacroix C., Monsemp C., Simon S.,
RA Harris B., Atkin R., Doggett J., Mayes R., Keating L., Wheeler P.R.,
RA Parkhill J., Barrell B.G., Cole S.T., Gordon S.V., Hewinson R.G.;
RT "The complete genome sequence of Mycobacterium bovis.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:7877-7882(2003).
DR EMBL; BX248342; CAD97414.1; -.
DR GO; GO:0008415; F:acyltransferase activity; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.

```

DR GO; GO:0016740; F:transferase activity; IEA.  
 DR GO; GO:0006633; P:fatty acid biosynthesis; IEA.  
 DR GO; GO:0008152; P:metabolism; IEA.  
 DR InterPro; IPR001227; Ac transferase.  
 DR InterPro; IPR000794; Ketoacyl synth.  
 DR InterPro; IPR002539; Maoc dehydratas.  
 DR Pfam; PF00698; Acyl transf 1; 1.  
 DR Pfam; PF00109; ketoacyl-synt; 1.  
 DR Pfam; PF02801; Ketoacyl-synt\_C; 1.  
 DR Pfam; PF01575; Maoc dehydratas; 1.  
 DR PROSITE; PS00606; B\_KETOACYL\_SYNTHASE; UNKNOWN\_1.  
 KW Acyltransferase; Complete proteome; Transferase.  
 SQ SEQUENCE 3069 AA; 326304 MW; 6855667940C025A2 CRC64;

Query Match 59.1%; Score 39; DB 2; Length 3069;  
 Best Local Similarity 80.0%; Pred. No. 1.5e+03;  
 Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 PFATELHERL 13  
 | | | | | | | |  
 Db 2881 PNETELHERL 2890

## RESULT 33

Q9X7E2 PRELIMINARY; PRT; 3076 AA.  
 AC Q9X7E2;  
 DT 01-NOV-1999 (TREMBLrel. 12, Created)  
 DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)  
 DT 25-OCT-2004 (TREMBLrel. 28, Last annotation update)  
 DE Fatty acid synthase.  
 GN Name-fas; OrderedLocusNames=ML1191;  
 OS Mycobacterium leprae.  
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
 OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.  
 OX NCBI\_TaxID=1769;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RK MEDLINE=93188700; PubMed=8446027;  
 RA Eiglmeyer K., Honore N., Woods S.A., Caudron B., Cole S.T.;  
 RT "Use of an ordered cosmid library to deduce the genomic organization  
 of Mycobacterium leprae.";  
 RL Mol. Microbiol. 7:197-206(1993).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Seeger K.J., Harris D.;  
 RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RA James K.D., Parkhill J., Barrell B.G., Rajandream M.A.;  
 RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.  
 RN [4]  
 RP SEQUENCE FROM N.A.

RC STRAIN=TN;  
 RX MEDLINE=21128732; PubMed=11234002; DOI=10.1038/35059006;  
 RA Cole S.T., Eiglmeyer K., Parkhill J., James K.D., Thomson N.R.,  
 RA Wheeler P.R., Honore N., Garnier T., Churcher C.M., Harris D.E.,  
 RA Mungall K.L., Basham D., Brown D., Chillingworth T., Connor R.,  
 RA Davies R.M., Devlin K., Duthoy S., Feltwell T., Fraser A., Hamlin N.,  
 RA Holroyd S., Hornsby T., Jagels K., Lacroix C., Maclean J., Moule S.,  
 RA Murphy L.D., Oliver K., Quail M.A., Rajandream M.A., Rutherford K.M.,  
 RA Rutter S., Seeger K., Simon S., Simmonds M., Skelton J., Squares R.,  
 RA Squares S., Stevens K., Taylor K., Whitehead S., Woodward J.R.,  
 RA Barrell B.G.;  
 RT "Massive gene decay in the leprosy bacillus.";  
 RL Nature 409:1007-1011(2001).  
 DR EMBL; AL049478; CAB39571.1; -;  
 DR EMBL; AL583921; CAC31572.1; -;  
 DR PIR; A87058; A87058.  
 DR HSSP; P25715; IMLA.  
 DR Leprowa; ML1191; -;  
 DR GO; GO:0005835; C:fatty-acid synthase complex; IEA.  
 DR GO; GO:0004312; F:fatty-acid synthase activity; IEA.

DR GO; GO:0016491; F:oxidoreductase activity; IEA.  
 DR GO; GO:0016740; F:transferase activity; IEA.  
 DR GO; GO:0006633; P:fatty acid biosynthesis; IEA.  
 DR GO; GO:0008152; P:metabolism; IEA.  
 DR InterPro; IPR001227; Ac transferase.  
 DR InterPro; IPR003965; Fatty acid synth.  
 DR InterPro; IPR000794; Ketoacyl synth.  
 DR InterPro; IPR002539; Maoc dehydratas.  
 DR Pfam; PF00698; Acyl transf 1; 1.  
 DR Pfam; PF00109; ketoacyl-synt; 1.  
 DR Pfam; PF02801; Ketoacyl-synt\_C; 1.  
 DR Pfam; PF01575; Maoc dehydratas; 1.  
 DR PRINTS; PR01483; FASYNTHASE.  
 DR PROSITE; PS00606; B\_KETOACYL\_SYNTHASE; UNKNOWN\_1.  
 KW Complete proteome; Transferase.  
 SQ SEQUENCE 3076 AA; 328358 MW; D88556730478E549 CRC64;

Query Match 59.1%; Score 39; DB 2; Length 3076;  
 Best Local Similarity 80.0%; Pred. No. 1.5e+03;  
 Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 PFATELHERL 13  
 | | | | | | | |  
 Db 2888 PNETELHERL 2897

## RESULT 34

Q73XH7 PRELIMINARY; PRT; 3092 AA.  
 ID Q73XH7;  
 AC Q73XH7;  
 DT 05-JUL-2004 (TREMBLrel. 27, Created)  
 DT 05-JUL-2004 (TREMBLrel. 27, Last sequence update)  
 DT 05-JUL-2004 (TREMBLrel. 27, Last annotation update)  
 DE Fas.  
 GN Name-fas; OrderedLocusNames=MAP2332c;  
 OS Mycobacterium paratuberculosis.  
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
 OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.  
 OX NCBI\_TaxID=1770;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=K10;  
 RA Li L., Bannantine J., Zhang Q., Anonsin A., Alt D., Kapur V.;  
 RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AE017235; AAS04649.1; -;  
 DR GO; GO:0005835; C:fatty-acid synthase complex; IEA.  
 DR GO; GO:0004312; F:fatty-acid synthase activity; IEA.  
 DR GO; GO:0016491; F:oxidoreductase activity; IEA.  
 DR GO; GO:0016740; F:transferase activity; IEA.  
 DR GO; GO:0006633; P:fatty acid biosynthesis; IEA.  
 DR GO; GO:0008152; P:metabolism; IEA.  
 DR InterPro; IPR001227; Ac transferase.  
 DR InterPro; IPR003965; Fatty acid synth.  
 DR InterPro; IPR002114; HPI\_SerP\_S.  
 DR InterPro; IPR000794; Ketoacyl synth.  
 DR InterPro; IPR002539; Maoc dehydratas.  
 DR Pfam; PF00698; Acyl transf 1; 1.  
 DR Pfam; PF00109; ketoacyl-synt; 1.  
 DR Pfam; PF02801; Ketoacyl-synt\_C; 1.  
 DR Pfam; PF01575; Maoc dehydratas; 1.  
 DR PRINTS; PR01483; FASYNTHASE.  
 DR PROSITE; PS00606; B\_KETOACYL\_SYNTHASE; UNKNOWN\_1.  
 DR PROSITE; PS00589; PTS\_HPR\_SER; UNKNOWN\_1.  
 KW Complete proteome; Transferase.  
 SQ SEQUENCE 3092 AA; 328070 MW; 3616EDD7DF0350D9 CRC64;

Query Match 59.1%; Score 39; DB 2; Length 3092;  
 Best Local Similarity 80.0%; Pred. No. 1.5e+03;  
 Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 PFATELHERL 13  
 | | | | | | | |  
 Db 2901 PNETELHERL 2910



Query Match 57.6%; Score 38; DB 2; Length 162;  
 Best Local Similarity 61.5%; Pred. No. 1.2e+02;  
 Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 KLVPPFATLHERL 13  
 DB 126 KIVPFALERETRL 138

## RESULT 37

Q6HCP9 PRELIMINARY; PRT; 176 AA.  
 ID Q6HCP9  
 AC Q6HCP9;  
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)  
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)  
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)  
 DE Hypothetical protein.  
 GN OrderedLocusNames=BF9727\_4362;  
 OS Bacillus thuringiensis (subsp. konkukian).  
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.  
 OX NCBI\_TaxID=180856;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=97-27;  
 RA Bretin T.S., Bruce D., Challacombe J.F., Gilna P., Han C., Hill K.,  
 RA Hitchcock P., Jackson P., Keim P., Longmire J., Lucas S., Okinaka R.,  
 RA Richardson P., Rubin E., Tice H.;  
 RT "Complete genome sequence of Bacillus thuringiensis 97-27.";  
 RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AE017355; AAT60913.1; -.  
 KW Complete proteome.  
 SQ SEQUENCE 176 AA; 20899 MW; C2BFA22CCB7F18A8 CRC64;

Query Match 57.6%; Score 38; DB 2; Length 176;  
 Best Local Similarity 54.5%; Pred. No. 1.3e+02;  
 Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 KLVPPFATLHE 11  
 DB 75 KNAPFSTEIHD 85

## RESULT 38

Q9X548 PRELIMINARY; PRT; 250 AA.  
 ID Q9X548  
 AC Q9X548; 1999 (TrEMBLrel. 12, Created)  
 DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE Yag1C.  
 OS Corynebacterium glutamicum (Brevibacterium flavum).  
 OG Plasmid R-plasmid PAG1.  
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
 OC Corynebacteriaceae; Corynebacteriaceae; Corynebacterium.  
 OX NCBI\_TaxID=1718;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=22243;  
 RX MEDLINE=20532766; PubMed=11078655; DOI=10.1006/plas.2000.1489;  
 RA Tauch A., Puhler A., Kallinowski J., Thierbach G.;  
 RT "TetZ, a new tetracycline resistance determinant discovered in gram-  
 positive bacteria, shows high homology to gram-negative regulated  
 efflux systems";  
 RL Plasmid 44:285-291 (2000).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=22243;  
 RA Tauch A.;  
 RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF121000; AAD25067.1; -.  
 KW Plasmid.  
 SQ SEQUENCE 250 AA; 27505 MW; 436F69A27AFA9DBF CRC64;

## RESULT 35

Q6K305 PRELIMINARY; PRT; 57 AA.  
 ID Q6K305  
 AC Q6K305;  
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)  
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)  
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)  
 DE Hypothetical protein OSJNB0066C12.39.  
 GN Names=OSJNB0066C12.39;  
 OS Oryza sativa (japonica cultivar-group).  
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC Ehrhartoideae; Oryzoae; Oryza.  
 OX NCBI\_TaxID=39947;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Sasaki T., Matsumoto T., Katayose Y.;  
 RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 9, BAC  
 RT clone:OSJNB0066C12.";  
 RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AP005738; BAD23565.1; -.  
 KW Hypothetical protein.  
 SQ SEQUENCE 57 AA; 6414 MW; 723678FE6D6828F6 CRC64;

Query Match 57.6%; Score 38; DB 2; Length 57;  
 Best Local Similarity 80.0%; Pred. No. 41;  
 Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 LVPFATLHE 11  
 DB 6 LVPFATLHE 15

## RESULT 36

Q89SR9 PRELIMINARY; PRT; 162 AA.  
 ID Q89SR9  
 AC Q89SR9;  
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
 DE Transcriptional regulatory protein.  
 GN OrderedLocusNames=blr2331;  
 OS Bradyrhizobium japonicum.  
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;  
 OC Bradyrhizobiaceae; Bradyrhizobium.  
 OX NCBI\_TaxID=375;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=USDA110;  
 RX MEDLINE=22484998; PubMed=12597275;  
 RA Kaneko T., Nakamura Y., Sato S., Minamisawa K., Uchiumi T.,  
 RA Sasamoto S., Watanabe A., Idesawa K., Iriiguchi M., Kawashima K.,  
 RA Kohara M., Matsumoto M., Shimpou S., Tsuruoka H., Wada T., Yamada M.,  
 RA Tabata S.;  
 RT "Complete genomic sequence of nitrogen-fixing symbiotic bacterium  
 Bradyrhizobium japonicum USDA110.";  
 RL DNA Res. 9:189-197 (2002).  
 CC -1- SIMILARITY: Contains 1 HTH marR-type DNA-binding domain.  
 DR EMBL; AP005943; BAC47596.1; -.  
 DR GO; GO:0005622; C:intracellular; IEA.  
 DR GO; GO:0003700; F:transcription factor activity; IEA.  
 DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.  
 DR InterPro; IPR000835; HTH\_MarR.  
 DR IPR000835; HTH\_MarR.  
 DR Pfam; PF01047; MarR.1  
 DR PRINTS; PR00598; HTHMARR.  
 DR SMART; SM00347; HTH\_MARR; 1.  
 KW Complete proteome; DNA-binding; Transcription; Transcription regulation.  
 SQ SEQUENCE 162 AA; 18163 MW; PAE5DAB4137F535A CRC64;

```
Query Match 57.6%; Score 38; DB 2; Length 250;
Best Local Similarity 87.5%; Pred. No. 1.8e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 PFATELHE 11
DB 202 PFADLHE 209

RESULT 39
ID O33624 PRELIMINARY; PRT; 288 AA.
AC O33624;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Orf5yn (Fragment).
OS Streptomyces coelicolor.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomycetes.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=M145;
RA Burger A., Wohlleben W.;
RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; Y14206; CAA74602.1;
DR InterPro; IPR000834; Peptidase_M14.
DR PROSITE; PS00133; CARBOXYPEPT_ZN_2; UNKNOWN_1.
FT NON_TER 288
SQ SEQUENCE 288 AA; 31935 MW; A20726737132F121 CRC64;

Query Match 57.6%; Score 38; DB 2; Length 288;
Best Local Similarity 58.3%; Pred. No. 2.1e+02;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 LVPFATELHERL 13
DB 273 LVPLAETVHERM 284

RESULT 40
Q6GK71
ID O6GK71 PRELIMINARY; PRT; 311 AA.
AC O6GK71;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Putative inosine-uridine preferring nucleoside hydrolase.
GN OrderedLocNames-SAR0236;
OS Staphylococcus aureus (strain MRSA252).
OC Bacteria; Firmicutes; Bacilliales; Staphylococcus.
OX NCBI_TaxID=282458;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=15213324; DOI=10.1073/pnas.04025211101;
RA Holden M.T.G., Fell E.J., Lindsey J.A., Peacock S.J., Day N.P.J.,
RA Enright M.C., Foster F.J., Moore C.E., Hurst L., Atkin R., Barron A.,
RA Bason N., Bentley S.D., Chillingworth C., Chillingworth T.,
RA Churcher C., Clark L., Corton C., Cronin A., Doggett J., Dowd L.,
RA Feltwell T., Hance Z., Harris B., Hauser H., Holroyd S., Jagels K.,
RA James K.D., Lennard N., Line A., Maves R., Moule S., Mungall K.,
RA Ormond D., Quail M.A., Rabinowitsch E., Rutherford K.M., Sanders M.,
RA Sharp S., Simmonds M., Stevens K., Whitehead S., Barrell B.G.,
RA Spratt B.G., Parkhill J.;
RT "Complete genomes of two clinical Staphylococcus aureus strains:
RT evidence for the rapid evolution of virulence and drug resistance.";
RL Proc. Natl. Acad. Sci. U.S.A. 101:9786-9791(2004).
DR EMBL; BX571856; CAG39263.1;
DR InterPro; IPR002048; EF-hand.
DR InterPro; IPR001910; I/U_nucled_hydro.
DR Pfam; PF01156; IU_nuc_hydro; 1.
DR ProDom; PD007736; I/U_nucled_hydro; 1.
```

---

```
DR PROSITE; PS00018; EF_HAND; UNKNOWN_1.
DR PROSITE; PS01247; IUNH; 1.
KW Complete proteome.
SQ SEQUENCE 311 AA; 33881 MW; FC8A3429B4309D6A CRC64;

Query Match 57.6%; Score 38; DB 2; Length 311;
Best Local Similarity 70.0%; Pred. No. 2.2e+02;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 KLVPFATELH 10
DB 74 KLVAFASEIH 83

RESULT 41
Q6RWS4
ID Q6RWS4 PRELIMINARY; PRT; 314 AA.
AC Q6RWS4;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Nitrilase (EC 3.5.5.7).
GN ORFNames=BD7612;
OS uncultured organism.
OX unclassified; environmental samples.
OX NCBI_TaxID=155900;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=15066841; DOI=10.1128/AEM.70.4.2429-2436.2004;
RA Robertson D.E., Chaplin J.A., DeSantis G., Fodor M., Madden M.,
RA Chi E., Richardson T., Milan A., Miller M., Weiner D.P., Wong K.,
RA McQuaid J., Farrell B., Preston L.A., Tan X., Sneed M.A., Keller M.,
RA Mathur E., Kretz P.L., Burk M.J., Short J.M.;
RT "Exploring nitrilase sequence space for enantioselective catalysis.";
RL Appl. Environ. Microbiol. 70:2429-2436(2004).
DR EMBL; AY487428; AAR97375.1;
DR GO; GO:0018762; F:aliphatic nitrilase activity; IEA.
DR GO; GO:0016810; F:hydrolase activity, acting on carbon-nitrog. .; IEA.
DR GO; GO:0006807; P:nitrogen metabolism; IEA.
DR InterPro; IPR000132; N/CN_hydrtase.
DR InterPro; IPR003010; Ntlse/CNhydrtase.
DR Pfam; PF00795; CN_hydrolase; 1.
DR PROSITE; PS00263; CN_HYDROLASE; 1.
DR PROSITE; PS00920; NITRIL_CHT_1; 1.
KW Hydrolase.
SQ SEQUENCE 314 AA; 34391 MW; B6C814543EC19B83 CRC64;

Query Match 57.6%; Score 38; DB 2; Length 314;
Best Local Similarity 72.7%; Pred. No. 2.3e+02;
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 VPFATELHERL 13
DB 63 VKLAELHERL 73

RESULT 42
O66538
ID O66538 PRELIMINARY; PRT; 317 AA.
AC O66538;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE ADP-heptose:LPS heptosyltransferase.
GN Name-rfaC2; OrderedLocNames=AQ_145;
OS Aquifex aeolicus.
OC Bacteria; Aquificae; Aquificales; Aquificaceae; Aquifex.
OX NCBI_TaxID=63363;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VF5;
RX MEDLINE=98196666; PubMed=9537320; DOI=10.1038/32831;
RA Deckert G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,
```

RA Graham D.E., Overbeek R., Sneed M.A., Keller M., Aujay M., Huber R.,  
RA Feldman R.A., Short J.M., Olsen G.J., Swanson R.V.,  
RT "The complete genome of the hyperthermophilic bacterium Aquifex  
RT aeolicus".  
RL Nature 392:353-358(1998).  
DR EMBL; A5000675; AAC06484.1; -.  
DR PIR; G70313; G70313.  
DR GO; GO:0016740; F:transferase activity; IEA.  
DR GO; GO:0008152; P:metabolism; IEA.  
DR InterPro; IPR002201; Glyco trans.  
DR Pfam; PF01075; Glyco\_transf\_9; 1.  
KW Complete proteome.  
SQ SEQUENCE 317 AA; 36675 MW; F5705F774129F35D CRC64;

Query Match 57.6%; Score 38; DB 2; Length 317;  
Best Local Similarity 72.7%; Pred. No. 2.3e+02;  
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 KLVPFATLHE 11  
|||||  
DB 110 KLVPFHWELHE 120

RESULT 43  
Q8TYAI PRELIMINARY; PRT; 319 AA.  
AC Q8TYAI;  
DT 01-JUN-2002 (TrEMBLrel. 21, Created)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE Ca2+/Na+ antiporter.  
GN Name=ecm27.1; OrderedLocusNames=MK0404;  
OS Methanopyrus kandleri.  
OC Archaea; Euryarchaeota; Methanopyri; Methanopyrales; Methanopyraceae;  
OC Methanopyrus.  
OX NCBI\_TaxID=2320;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=AV19 / DSM 6324 / JCM 9639;  
RX MEDLINE=2197647; PubMed=11930014; DOI=10.1073/pnas.032671499;  
RA Slesarev A.I., Mezheva K.V., Makarova K.S., Polushin N.N.,  
RA Shcherbinina O.V., Shakhova V.V., Belova G.I., Aravind L.,  
RA Natale D.A., Rogozin I.B., Tatusov R.L., Wolf Y.I., Stetter K.O.,  
RA Malykh A.G., Koonin E.V., Kozlyavkin S.A.;  
RT "The complete genome of hyperthermophile Methanopyrus kandleri AV19  
RT and monophyly of archaeal methanogens".  
RL Proc. Natl. Acad. Sci. U.S.A. 99:4644-4649(2002).  
DR EMBL; AE010337; AM01619.1; -.  
DR GO; GO:0016021; C:integral to membrane; IEA.  
DR Pfam; PF01699; Na\_Ca\_ex; 2.  
KW Complete proteome.  
SQ SEQUENCE 319 AA; 34132 MW; 1B17C12E02A1030C CRC64;

Query Match 57.6%; Score 38; DB 2; Length 319;  
Best Local Similarity 66.7%; Pred. No. 2.3e+02;  
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 LVPFATLHERL 13  
|||||  
DB 221 LAPFATLPEKL 232

RESULT 44  
Q67L63 PRELIMINARY; PRT; 330 AA.  
AC Q67L63;  
DT 25-OCT-2004 (TrEMBLrel. 28, Created)  
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)  
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)  
DE Aldehyde ferredoxin oxidoreductase.  
GN ORFNames=STH2598;  
OS Symbiobacterium thermophilum.  
OC Bacteria; Actinobacteria; Symbiobacterium.

OX NCBI\_TaxID=2734;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=IAM14863;  
RA Ueda K., Yamashita A., Ishikawa J., Shimada M., Watsui T.,  
RA Morikura K., Ikeda H., Hattori M., Beppu T.;  
RT "Complete genome sequence of an uncultured bacterium Symbiobacterium  
RT thermophilum".  
RL Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF006840; BAD41583.1; -.  
DR InterPro; IPR001203; Oxred\_Ald\_Fedxn.  
DR Pfam; PF02730; AFOR\_N; 1.  
SQ SEQUENCE 330 AA; 34622 MW; 359B740518704F9D CRC64;

Query Match 57.6%; Score 38; DB 2; Length 330;  
Best Local Similarity 77.8%; Pred. No. 2.4e+02;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 VPFATLHE 11  
|||||  
DB 158 VPFATWVHE 166

RESULT 45  
Q6R3I5 PRELIMINARY; PRT; 335 AA.  
AC Q6R3I5;  
DT 05-JUL-2004 (TrEMBLrel. 27, Created)  
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)  
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)  
DE Conserved protein.  
GN Gemmata sp. Wal-1.  
OS Bacteria; Planctomycetes; Planctomycetacia; Planctomycetales;  
OC Planctomycetaceae; Gemmata.  
OX NCBI\_TaxID=235140;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX PubMed=15014146; DOI=10.1093/molbev/meh113;  
RA Chistoserdova L., Jenkins C., Kalyuzhnyaya M.G., Marx C.J., Lapidus A.,  
RA Vorholt J.A., Staley J.T., Lidstrom M.E.;  
RT "The enigmatic planctomycetes may hold a key to the origins of  
RT methanogenesis and methylotrophy".  
RL Mol. Biol. Evol. 21:1234-1241(2004).  
DR EMBL; AY515680; AAS88969.1; -.  
DR GO; GO:0016787; F:hydrolase activity; IEA.  
DR InterPro; IPR002756; DUF66.  
DR InterPro; IPR002821; Hydantoinase\_A.  
DR Pfam; PF01968; Hydantoinase\_A; 1.  
DR ProDom; PD024636; DUF66; 1.  
SQ SEQUENCE 335 AA; 36242 MW; D0094C3791B10372 CRC64;

Query Match 57.6%; Score 38; DB 2; Length 335;  
Best Local Similarity 46.2%; Pred. No. 2.4e+02;  
Matches 6; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 KLVPFATLHERL 13  
::: |||: |||  
DB 244 QIIHFATRIHSRL 256

Search completed: September 1, 2005, 16:40:43  
Job time : 178 secs

**This Page Blank (uspto)**